

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 CompuGen Ltd.

# OM protein - nucleic search, using frame\_plus\_p2n model

Run on: August 17, 2003, 18:47:50 ; Search time 1902.81 Seconds  
(without alignments)  
2644.453 Million cell updates/sec

Title: US-09-786-715-8

Perfect score: 618  
Sequence: 1 MAVEBEGVIGVHTVDEWKL.....DDLQATIAKHAASVAASS 123

## Scoring table:

BLOSUM62	
Xgapop 10.0 , Xgapext 0.5	
Ygapop 10.0 , Ygapext 0.5	
Fgapop 6.0 , Fgapext 7.0	
Delop 6.0 , Delext 7.0	

Searched: 2888711 seqs, 2045481386 residues

Total number of hits satisfying chosen parameters: 5777422

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

## Command line parameters:

-MODE=frame+pn.model -DEV=slp  
-Q/cg2.1/USPTO.spool/US09786715.r/unat\_11082003.150514.6035/app\_query.fasta\_1.1052  
-DB=GenEmbl -QFWT=fastap -SUFFIX=rgc -MINMATCH=0.1 -LOOPEXT=0 -LOOPEXT=0  
-UNITS=bits -START=1 -END=1 -MATRIX=blomsu62 -TRANS=human40.cdi -LIST=45  
-DOCALIGN=200 -THR.SCOR=pcr -THR.MAX=100 -THR.MIN=0 -ALIGN=15 -MODE=LOCAL  
-OUTFMT=ptc -NORF=ext -HEA=SIZE=500 -MINLEN=0 -MAXLEN=200000000  
-USER=US09786715.ecgn\_1.15721\_etunat\_11082003.150514.6035 -NCPU=6 -ICPU=3  
-NO\_MMAP -LARGEQUERY -NEG.SCOR=0 -WAIT -DSPBLOCK=100 -LONGLOG  
-DEV\_TIMEOUT=120 -WARN\_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6  
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

## Database :

GenEmbl:\*  
1: gb\_da:\*  
2: gb\_htg:\*  
3: gb\_in:\*  
4: gb\_om:\*  
5: gb\_ov:\*  
6: gb\_pat:\*  
7: gb\_ph:\*  
8: gb\_pl:\*  
9: gb\_pr:\*  
10: gb\_ro:\*  
11: gb\_sts:\*  
12: gb\_sy:\*  
13: gb\_un:\*  
14: gb\_vl:\*  
15: em\_ba:\*  
16: em\_fun:\*  
17: em\_hum:\*  
18: em\_in:\*  
19: em\_mu:\*  
20: em\_om:\*  
21: em\_or:\*  
22: em\_ov:\*  
23: em\_pat:\*  
24: em\_ph:\*  
25: em\_pl:\*  
26: em\_ro:\*  
27: em\_sts:\*  
28: em\_un:\*

29: em\_vi:\*  
30: em\_htg\_hum:\*  
31: em\_htg\_inv:\*  
32: em\_htg\_other:\*  
33: em\_htg\_mus:\*  
34: em\_htg\_pln:\*  
35: em\_htg\_rod:\*  
36: em\_htg\_mam:\*  
37: em\_htg\_vrt:\*  
38: em\_sy:\*  
39: em\_htgo\_hum:\*  
40: em\_htgo\_mus:\*  
41: em\_htgo\_other:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	444	71.8	664	8	AF323593	AF323593 Prunus pe
2	442	71.5	665	8	AY170650	AY170650 Pisum sat
3	425	68.8	593	8	PSA19808	AJ319808 Pisum sat
4	423.5	68.5	603	8	RC7H10RXN	Z70677 R.communis
5	410	66.3	345	6	AX505468	AX505468 Sequence
6	410	66.3	480	8	AY088687	AY088687 Arabidops
7	410	66.3	497	8	ATTH10ARA	Z14084 A.thaliana
8	406	65.7	526	11	G73679	G73679 R2488R etio
9	406	65.7	601	8	OS092541	U92541 Oryza sativ
10	406	65.7	686	6	E08194	E08194 RPS13 gene
11	406	65.7	687	8	RIC7H	D21836 Oryza sativ
12	404	65.4	698	8	NTTRNA	X58537 N.tabacum m
13	401	64.9	345	8	AF483265	AF483265 Populus t
14	397	64.2	596	6	TA89762	AJ009762 Trilicium
15	393	63.6	784	8	AY271308	AY271308 Citrus x
16	392	63.4	393	6	A48516	A48516 Sequence 4
17	392	63.4	437	8	AF420472	AF420472 Trilicium
18	392	63.4	630	6	A48520	A48520 Sequence 8
19	392	63.4	630	8	TDAJ1903	AJ001903 Trilicium
20	392	63.4	653	6	AR016869	AR016869 Sequence
21	392	63.4	653	6	AR020895	AR020895 Sequence
22	392	63.4	653	6	AR027218	AR027218 Sequence
23	392	63.4	653	6	AR038505	AR038505 Sequence
24	392	63.4	653	6	AR064647	AR064647 Sequence
25	392	63.4	653	6	AR067572	AR067572 Sequence
26	392	63.4	653	6	I38524	I38524 Sequence 10
27	392	63.4	653	6	I56899	I56899 Sequence 10
28	392	63.4	653	6	I59665	I59665 Sequence 10
29	392	63.4	653	6	I75192	I75192 Sequence 10
30	388.5	62.9	580	8	BNU59380	U59380 Brassica na
31	388	62.8	384	6	A48514	A48514 Sequence 2
32	388	62.8	659	6	A48519	A48519 Sequence 7
33	388	62.8	670	6	TATH10RDH	X69915 T.aestivum
34	383	62.0	366	6	AX654096	AX654096 Sequence
35	383	62.0	740	8	ABO53294	ABO53294 Oryza sat
36	380.5	61.6	360	6	AX505469	AX505469 Sequence
37	380.5	61.6	360	8	BT004710	BT004710 Arabidops
38	380.5	61.6	366	8	AK118542	AK118542 Arabidops
39	378	61.2	353	8	AF286593	AF286593 Trilicium
40	378	61.2	629	8	TA804845	AJ310990 Pisum sat
41	378	61.2	630	8	PSA310990	AJ310990 Pisum sat
42	377.5	61.1	642	8	ATTH10RD1	Z35473 A.thaliana
43	376	60.8	357	8	AY040028	AY040028 Arabidops
44	376	60.8	328	8	ATTH10RD4	Z35476 A.thaliana
45	376	60.8	530	8	AK118035	AK118035 Arabidops

## ALIGNMENTS

RESULT 1

AF323593	664 bp	mRNA	linear	PLN 02-NOV-2001
LOCUS	AF323593			
DEFINITION	Prunus persica thiodoxin H mRNA, complete cds.			
ACCESSION	AF323593			
VERSION	AF323593.1			
KEYWORDS	GI:16588842			
SOURCE	Prunus persica (peach)			
ORGANISM	Prunus persica			
REFERENCE	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids; eurosids I; Rosales; Rosaceae; Amygdaloideae; Prunus.			
AUTHORS	1 (bases 1 to 664)			
TITLE	Callahan A.M., Morgens, P.H. and Cohen, R.A. Isolation and initial characterization of cDNAs for mRNAs regulated during peach fruit development			
JOURNAL	J. Am. Soc. Hortic. Sci. 118, 531-537 (1993)			
REFERENCE	2 (bases 1 to 664)			
AUTHORS	Callahan, A.M., Morgens, P.H., Cohen, R.A. and Scorza, R.			
TITLE	Regulation of peach gene expression in a peach/almond hybrid			
JOURNAL	Unpublished			
REFERENCE	3 (bases 1 to 664)			
AUTHORS	Callahan, A.M., Dunn, L.L. and Cohen, R.A.			
TITLE	Direct Submission			
JOURNAL	Submitted (22-NOV-2000) USDA-ARS, Appalachian Fruit Research Station, 45 Wiltshire Rd., Kearneysville, WV 25430, USA			
FEATURES	Location/Qualifiers			
Source	1..664			
	/organism="Prunus persica"			
	/mol_type="mRNA"			
	/culivar="B612615"			
	/db_xref="taxon:3760"			
	/clone="pch306c"			
	/tissue_type="ripe fruit"			
	71..481			
	/codon_start=1			
	/product="Chlorodoxin H"			
	/protein_id="AA126915.1"			
	/db_xref="GI:16588843"			
	/translation="MAENQYIGCHTTOAEWEOHLKGNENKLVVDFTASMCPCRLIAPLIELAKKTPVETLKDVDLRTLSEBQVEAMPTEFLKEGIVDKVVGAKDD ELQIKYAKHVAANAASATSAATAAATATASA"			
CDS				
	182 a 147 c 161 g 174 t			
BASE COUNT				
ORIGIN				
Alignment Scores:				
Pred. No.:	1.27e-41	Length:	664	
Score:	444.00	Matches:	81	
Percent Similarity:	84.03%	Conservative:	19	
Best Local Similarity:	68.07%	Mismatches:	19	
Query Match:	71.84%	Indels:	0	
DB:	8	Gaps:	0	
US-09-786-715-8 (1-123) x AF323593 (1-664)				
OY	5	GLUGLUGLUGLGNVALIIIEGLYVALHISHTYVALASPLGUTRTPYLSLEGNLLEUGLNASN	24	
DB	77	GAGAAATATCAAGTCATCGCTGCCACACTCACTCAAGCCGTGGAAAGACGACTCCATAAG	136	
OY	25	AlalysAsperLysLysLeuLleValValAspPheThrAlaSerTPCysGlyProCys	44	
DB	137	GGAACGAGACAAAGAAATCGTGGTGGATTTTCACGCTTCTGTGGTGACCGTGC	196	
OY	45	ArgPheMetAlaProValLeuAlaGluLleAlaLysLysThrProGluLeuLlePheLeu	64	
DB	197	CGGTTGATCGCCCAATCTTGCGGAGTGGCTTAAGAACCCAGAACGACTTCCCTA	256	
OY	65	LysValAspValAspGluValArgProValAlaGluGluTyrSerLleGluAlaMetPro	84	
DB	257	AAGCGGACGCTGATGAACTGAGACACTTTCCAGAGAGTGGGTGGAGGCAATGCTCT	316	
OY	85	ThrPheLeuPheLeuLysAspClyGluLleValAspLysValValGlyAlaSerLysAsp	104	
DB	317	ACCTTCCCTCTTCACGAGAGGCAAGATATGTGACAAAGTTGTGGTCCCAAGAAAGAC	376	

OY		105	Asplenginalatnrllealalyhislaserlavalalaialalaserser	123
Dd	377	GAGTTACAGATCAAAAGTTGGCCAGACATGTCGCCTGCCGTGCCTCGCAC	433	
RESULT 2				
LOCUS	AY170650		665 bp	mRNA linear PLN 02-JAN-2003
DEFINITION	Pisum sativum clone 3 thiorodoxin h mRNA, complete cds.			
ACCESSION	AY170650			
VERSION	AY170650.1	GI:27466893		
KEYWORDS				
SOURCE				
ORGANISM				
	Pisum sativum (pea)			
	Pisum sativum			
	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;			
	Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;			
	rosids; eurosids I; Fabales; Fabaceae; Papilionoideae; Viciae;			
	Pisum.			
REFERENCE	1 (bases 1 to 665)			
AUTHORS	Morlichard,F., Renaud,M., Duval,F.D. and Machereel,D.			
TITLE	Expression of the NADP-thiorodoxin reductase/thiorodoxins h system during germination of seeds of Pisum sativum L			
JOURNAL	Unpublished			
REFERENCE	2 (bases 1 to 665)			
AUTHORS	Morlichard,F., Renaud,M., Duval,F.D. and Machereel,D.			
TITLE	Direct Submission			
JOURNAL	Submitted (30-OCT-2002) UMR 1191 Physiologie Molculaire des Semences, UFR Sciences, 16 bd Lavoisier, Angers 49045, France			
FEATURES				
source	Location/Qualifiers			
	1..665			
	/organism="Pisum sativum"			
	/mol_type="mRNA"			
	/db_xref="taxon:3888"			
	/clone="3"			
CDS	74..415			
	/note="trx h"			
	/codon_start=1			
	/product="thiorodoxin h"			
	/protein_id="AA012854.1"			
	/db_xref="GI:27466894"			
	/translation="MAEEGVYIGHYHTVDANKKEOLEKASKKLIVDFTSKSGPCRF IAPLPIELAKRLTHVTLKVDLVSEWGLEAMPFLPLKDGLVDKVGAKE ELOLRIDRHNAA"			
BASE COUNT	220 a 85 c 164 g 196 t			
ORIGIN				
Alignment Scores:				
Pred. No.:	2.15e-41	Length:	665	
Score:	442.00	Matches:	84	
Percent Similarity:	86.49%	Conservative:	12	
Best Local Similarity:	75.68%	Mismatches:	15	
Query Match:	71.52%	Indels:	0	
DB:	8	Gaps:	0	
US-09-786-715-8 (1-123) x AY170650 (1-665)				
OY	5	GlucIugIugiInValIIleglyValIHIsThrValAspGIuTrPrLyLSLeuGlnLeuGlnasn	24	
Dd	80	GAAGAGGACAGAAGTGCgTGTCACACAGCCTGATGCTTGGAAGAAAGATTAGAAAG	139	
OY	25	AlAlysaspsrLySLySLeuIlEvaIValAspPherThrAlaserTrpCySGlyProCys	44	
Dd	140	GGAAGAACGCTCAAGAAACATGATGTAGTTGATTTCACCTGCTTGTCGGTCATGC	199	
OY	45	ArgphemeIarProValIleuAlagIuIlIalalySyrHrrProGIuLeuIllepheIeu	64	
Dd	200	CGTTTATATGGCCCAATTTTGGCAGAAATGGCTATAAAGCTTACACATGTACTTTCCTT	259	
OY	65	LysValaspValAspGIuValArpProValIalagIuIlurSerIllegIuaIMetPro	84	
Dd	260	AAGGTTGACGGTAGTAATTAACAAGACTGTTCCAGAGAGTGGGAATTTGAAGCATAGCCA	319	
OY	85	ThrheIeuPhelauIeulySaspGIyGIuIlIevaIalaspLySValValGIalaserLyasp	104	

```

DB      320 ACATCTGCTCTTGAAGATGTCGATGTCGACCAAGTGTGCGGCCAAGAGAG 379
        |||
QY      105 Aspleuginalathrilialalyshtsalaser 115
        |||
DB      380 GAGCTGCAATGAAATTGACAGCATGCGAGCT 412
        |||

RESULT 3
PSA319808
LOCUS   PSA319808          593 bp    mRNA    linear    PLN 17-JUN-2001
DEFINITION
Pisum sativum mRNA for thiorodoxin h (trxx gene).
ACCESSION
AJ319808.1 GI:14485508
VERSION
thiorodoxin h: trxx gene.
KEYWORDS
Pisum sativum (pea)
SOURCE
Pisum sativum
ORGANISM
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
rosids; eurosids I; Fabales; Fabaceae; Papilionoideae; Viciae;
Pisum.

REFERENCE
1 Navarro,E., Gomez,J.M., Jimenez,A. and Sevilla,F.
  TITLE
  Isolation of a new thiorodoxin h in pea plants. Expression during
  JOURNAL
  development and under oxidative stress induced conditions
  REFERENCE
  Unpublished
  AUTHORS
  2 (bases 1 to 593)
  TITLE
  Navarro,E.
  JOURNAL
  Direct Submission
  REFERENCE
  Submitted (13-JUN-2001) Navarro E., Nutrition and Plant Physiology,
  JOURNAL
  Cebas-csic, Cebas, Campus de Espinardo,, Murcia, 30100, SPAIN

FEATURES
location/Qualifiers
source
1..593
/organism="Pisum sativum"
/mol_type="mRNA"
/db_xref="taxon:3888"
/tissue_type="leaves"
1..593
/gene="trxx"
60..416
/gene="trxx"
/codon_start=1
/product="thiorodoxin h"
/protein_id="CAC42084.1"
/db_xref="GI:14485508"
/translation="MAENEVIAHSDSEKQIOKTESKKLIIVDFTSKCGPCRF
IAPILAEIAKRPQVIFLKVDELSEVAEEMSIEMPTFLLEKGMEDVYVGAKE
ELQIAITKHATTAVTA"

BASE COUNT      200 a      97 c      136 g      160 t
ORIGIN
Alignment Scores:
Pred. No.:      1.65e-39      Length:      593
Score:          425.00      Matches:      80
Percent Similarity: 83.48%      Conservative: 16
Best Local Similarity: 69.57%      Mismatches: 19
Query Match:    68.77%      Indels:      0
Gaps:           0

US-09-786-715-8 (1-123) x PSA319808 (1-593)
QY      6 GIUGIUGIValIleGlyValHisThrValAspGluTrpLysLeuGlnLeuGlnAsn 25
        |||
DB      69 GAAATGAGGTGATCGCTTCATTCGATGATGATGAGAGAGACAGATCCAGAAAGCA 128
        |||
QY      26 LysAspSerLysLysLeuIleValValAspPheThrAlaSerTrpCysGlyProCysArg 45
        |||
DB      129 ACCGATTCGCAAAAACTGATGTGTGATTTTACTGCTGCTGCGTGCAGCCGCT 188
        |||
QY      46 PheMetAlaProValLeuAlaGluIleAlaLysLysThrProGluLeuIlePheLeuLys 65
        |||
DB      189 TTTATTGCTCAATTCCTVAGCAGAGATGCTTAAGAGACACCTCAAGTATCTTCTCAAG 248
        |||
QY      66 ValAspValAspGluValArgProValAlaGluGluIuTrpSerIleGluAlaMetProThr 85

```

```

DB      249 GTGACACATTGACGATTTAGAGAGCTGTTCGCCAGGAATGCTATGAGAGCTATGCCAAC 308
        |||
QY      86 PheLeuPheLeuLysAspGlyGluIleValAspLysValValGlyAlaSerLysAspAsp 105
        |||
DB      309 TTCTGCTCTTGAAGAGAGCATGGAAGGCTAAGTAGTGTGCTGTAAGAGAGAGAA 368
        |||
QY      106 LeuGlnAlaThrIleAlaLysHisAlaSerAlaValAlaAlaAla 120
        |||
DB      369 CTGACGCTGGCAATACCAAGCATGACACACTGTTGCTACTGCT 413
        |||

RESULT 4
RCHTHORXN
LOCUS   RCHTHORXN          603 bp    mRNA    linear    PLN 04-APR-1996
DEFINITION
R. communis mRNA for thiorodoxin.
ACCESSION
270677
VERSION
270677.1 GI:1255953
KEYWORDS
thiorodoxin.
SOURCE
Ricinus communis (castor bean)
ORGANISM
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
rosids; eurosids I; Malpighiales; Euphorbiaceae; Ricinus.

REFERENCE
1 Szederkényi,J., Dolgenier,E. and Schobert,C.
  TITLE
  cDNA expressed in Ricinus cotyledons
  JOURNAL
  Unpublished
  REFERENCE
  2 (bases 1 to 603)
  TITLE
  Dolgenier,E.
  JOURNAL
  Direct Submission
  REFERENCE
  Submitted (04-APR-1996) Dolgenier E., University of Bayreuth,
  JOURNAL
  Department of Plant Physiology, Universitaetsstr.30, Bayreuth,
  Germany, 95440

FEATURES
location/Qualifiers
source
1..603
/organism="Ricinus communis"
/mol_type="mRNA"
/cultivar="Sanguineus"
/db_xref="taxon:3988"
/clone="PEDRH018"
/tissue_type="cotyledon"
/clone_lib="lambda Ecxell"
/dev_stage="seedling"
23..379
/codon_start=1
/product="thiorodoxin"
/protein_id="CAA94534.1"
/db_xref="GI:1255954"
/db_xref="SPTREMBL:Q43636"
/translation="MAEEGVIGCHTVEANNEQNDTKGLIVDFTSKCGPCR
FIAPFLAEIAKRLPVTFLKVDELKTVAHMVAESMPTFLKEGRIMDKVYGAKK
DELOOTIHKHATYST"

BASE COUNT      179 a      101 c      157 g      166 t
ORIGIN
Alignment Scores:
Pred. No.:      2.49e-39      Length:      603
Score:          423.50      Matches:      80
Percent Similarity: 83.05%      Conservative: 18
Best Local Similarity: 67.80%      Mismatches: 17
Query Match:    68.53%      Indels:      3
Gaps:           1

US-09-786-715-8 (1-123) x RCHTHORXN (1-603)
QY      5 GIUGIUGIValIleGlyValHisThrValAspGluTrpLysLeuGlnLeuGlnAsn 24
        |||
DB      32 GAAGAAGGCAAGATGATGCGGTGCCACACTGTGAGCATGATGACATATGCAGAG 91
        |||
QY      25 AlaLysAspSerLysLysLeuIleValValAspPheThrAlaSerTrpCysGlyProCys 44
        |||
DB      92 GAAATGATVACCAAGAGCATGCTGCTGATTTTACTGCTTCATGCTGAGTGCAGCATGC 151

```

FEATURES	source	Location/Qualifiers
1..345		
/organism="Arabidopsis thaliana"		
/mol_type="genomic DNA"		
/db_xref="taxon:3702"		
BASE COUNT	90 a	68 c 90 g 97 t
ORIGIN		
Alignment Scores:		
Pred. No.:	4.47e-38	Length: 345
Score:	410.00	Matches: 73
Percent Similarity:	85.32%	Conservative: 20
Best Local Similarity:	66.97%	Mismatch: 16
Query Match:	66.34%	Indels: 0
DB:	6	Gaps: 0
US-09-786-715-8 (1-133) x AX505468 (1-345)		
OY	5	gluGlugluGluValIleGlyValIHisThrValAspGluTrpLysLeuGluIleuGluAsn 24
Db	10	GAAAGAGACACAGATGATCGCTGCCACACCGCTTAGACATGAGACAGACGACGCTTCAAG 69
OY	25	AlaLysAspSerLysLysLeuIleValValAspPheThrAlaSerTrpCysGlyProCys 44
Db	70	GCTAATGATGCCAAACTCTTGCTGGTGATTTACACGCTTCTTGCTGGTGAGCACTGT 129
OY	45	ArgPheMetAlaProValLeuAlaGluIleAlaLysLysThrProGluLeuIlePheLeu 64
Db	130	CGTTTCATCCGCTCATTTCTTGGATTTGGCTTAAGAACTCTTCTTACGCTTTCCTC 189
OY	65	LysValAspValAspGluValArgProValAlaGluLysLysThrProGluLeuIlePheLeu 84
Db	190	AAGGTTATATCATGTGATGAATGAACTCGGTGCCAAGTATGGCGGATACAGGCATGCCA 249
OY	85	ThrPheLeuPheLeuLysAspGlyGluIleValAspLysValValGlyAlaSerLysAsp 104
Db	250	ACCTTCATGTTTTTGAAGGAGGAGAAATTTTGGACAAAGTTGTTGAGCCACGAAGAT 309

QY	105	Aspleniglatrnllealalyahs 113
Db	310	GAGCTTCAGTCTACCATTTGCCAACAC 336
RESULT 6		
AY088687		
LOCUS		480 bp mRNA linear PLN 14-APR-2003
DEFINITION	Arabidopsis thaliana clone 9189 mRNA, complete sequence.	
ACCESSION	AY088687	
VERSION	AY088687.1	GI:21407461
KEYWORDS	FLI_CDNA.	
SOURCE	Arabidopsis thaliana (thale cress)	
ORGANISM	Arabidopsis thaliana	
REFERENCE	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis. 1 (bases 1 to 480)	
AUTHORS	Haas,B.J., Volfovsky,N., Town,C.D., Troukhan,M., Alexandrov,N., Feldmann,K.A., Flavell,R.B., White,O. and Salzberg,S.L.	
TITLE	Full-length messenger RNA sequences greatly improve genome annotation	
JOURNAL	Genome Biol. 3 (6), RESEARCH0029 (2002)	
MEDLINE	22088475	
PUBMED	12093376	
REFERENCE	2 (bases 1 to 480)	
AUTHORS	Brover,V., Troukhan,M., Alexandrov,N., Lu,Y.-P., Flavell,R. and Feldmann,K.	
TITLE	Full-length cDNA from Arabidopsis thaliana	
JOURNAL	Unpublished	
REFERENCE	3 (bases 1 to 480)	
AUTHORS	Brover,V., Troukhan,M., Alexandrov,N., Lu,Y.-P., Flavell,R. and Feldmann,K.	
TITLE	Direct Submission	
JOURNAL	Submitted (11-MAR-2002) Ceres, Inc, 3007 Malibu Canyon Road, Malibu, CA 90265, USA	
COMMENT	This clone sequence is one of 5,000 Ceres full-length cDNAs made available to TIGR and Genbank. The following quality assessment of this set was done by comparison with known proteins: two percent of the clones are estimated to be 5'-truncated; less than one percent are 3'-truncated; approximately two percent represent alternative splice variants, including unspliced introns and spliced exons; one percent may contain premature stop codons; five percent may have frame shifts in a coding region. A sequence is considered to be 5'-truncated if it lacks the translation initiation start (ATG). A sequence is considered to be 3'-truncated if it lacks the C-terminal end of the encoded protein. Please note that these cDNA sequences are derived from the Ws or Luer ecotypes and therefore may contain polymorphisms when compared to sequences from Col-0. Genset carried out the library production and sequencing of the full-length clones. Ceres, Inc. carried out the clustering of the 5' sequences, selection of clones, and sequence assembly.	
FEATURES	Location/Qualifiers	
SOURCE	1..480	
	/organism="Arabidopsis thaliana"	
	/mol_type="mRNA"	
	/db_xref="taxon:3702"	
	/clone="9189"	
	13..357	
	/codon_start=1	
	/product="thioredoxin h"	
	/protein_id="AA067008.1"	
	/db_xref="GI:21617958"	
	/translation="MASEEGVAVACHTVETWNEQLOKANEKTLVVVDFTASGCPGR FIAFFEDLAKKLPNVFLKVDVDELKSVASDAVIAQMPFMFLKEGIDIKVYGAKK DELQSTAKHLA"	
BASE COUNT	124 a 93 c 111 g 152 t	
ORIGIN		
Alignment Scores:		
Pred. NO.:	6.64e-38	Length: 480
Score:	410.00	Matches: 73
Percent Similarity:	85.32%	Conservative: 20
Best Local Similarity:	66.97%	Mismatches: 16



Query Match: 66.34% Indels: 0  
 DB: 8 Gaps: 0  
 US-09-786-715-8 (1-123) x AT088687 (1-480)

QY 5 GluGluglyGlnValIleGlyValHsthRvalAspGluTrpLysLeuGlnAsn 24  
 |||||  
 DB 22 GAAGAAGGACAGTATGATCCCTGCCACACCGTTGACATGAGACAGACGCTTCAAG 81  
 QY 25 AlAlysAspSerLysLysLeuIleValAlaAspPheThAlaSerTrpCysGlyProCys 44  
 |||||  
 DB 82 GCTAATGAATCCAAACCTCTTGCGTGTGATTTCACGCGCTTCTTGCGTGTGACCATGT 141  
 QY 45 ArgPheMetAlaProValLeuAlaGluIleAlaLysLysThProGluLeuIlePheLeu 64  
 |||||  
 DB 142 GCTTTCATCGCTGCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 201  
 QY 65 LysValAspValAlaAspGluValArgProValAlaGluGluTrpSerIleGluAlaMetPro 84  
 |||||  
 DB 202 AAGGTTGATGACTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 261  
 QY 85 ThrPheLeuPheLeuLysAspGlyGluIleValAlaAspLysValAlaGlyAlaSerLysAsp 104  
 |||||  
 DB 262 ACCCTTCATGCTTTTGAAGGAGGAAGATTGACAAAGTTGAGACCAAGAAAGAT 321  
 QY 105 AspleuGlnAlaThrIleAlaLysHs 113  
 |||||  
 DB 322 GAGCTTCAGCTACCATTCACCAACAC 348

RESULT 7  
 ATTHIOARA 497 bp mRNA linear PLN 16-JUN-1994  
 LOCUS A.thaliana mRNA for thioredoxin H.  
 DEFINITION Z14084.1 GI:16551  
 VERSION Z14084.1 GI:16551  
 KEYWORDS thioredoxin H.  
 SOURCE Arabidopsis thaliana (thale cress)  
 ORGANISM Arabidopsis thaliana  
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
 rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsids.  
 1 (bases 1 to 497)  
 Riveria-Madrid,R., Martinho,P., Brugidou,C., Chartier,Y. and Meyer,Y.  
 Nucleotide sequence of a cDNA clone encoding an Arabidopsis  
 thaliana thioredoxin h

JOURNAL Plant Physiol. 102 (1), 327-328 (1993)  
 MEDLINE 94151431  
 PUBMED 8108503  
 REFERENCE 2 (bases 1 to 497)  
 Meyer Y.  
 TITLE Direct Submission  
 AUTHORS Submitted (16-JUN-1992) Yves Meyer PhD, Laboratoire de Physiologie  
 Journal Title vegetale, Universite de Perpignan, Av de Villeneuve 52, Perpignan,  
 66860, FRANCE

FEATURES  
 source Location/Qualifiers  
 1..497  
 /organism="Arabidopsis thaliana"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:3702"  
 30..374  
 /codon\_start=1  
 /product="thioredoxin H"  
 /protein\_id="CAA78462.1"  
 /db\_xref="GI:16551"  
 /db\_xref="SWISS-PROT:P29448"  
 /translation="MASEGQVIACHTEVFNNEQLOKANESKTLVVDFTSWCGPCR  
 FIAPFADIAKKLPVLYFLKVDDELKSVADWAIQAMPFTNLEKGLDKVYGAKK  
 DELOSTIKKHLA"

BASE COUNT 135 a 94 c 116 g 152 t  
 ORIGIN

Alignment Scores: 6..92e-38 Length: 497  
 Pred. No.:

Score: 410.00  
 Percent Similarity: 85.32%  
 Best Local Similarity: 66.97%  
 Query Match: 66.34%  
 DB: 8 Gaps: 0  
 US-09-786-715-8 (1-123) x ATTHIOARA (1-497)

QY 5 GluGluglyGlnValIleGlyValHsthRvalAspGluTrpLysLeuGlnAsn 24  
 |||||  
 DB 39 GAAGAAGGACAGTATGATCCCTGCCACACCGTTGACATGAGACAGACGCTTCAAG 98  
 QY 25 AlAlysAspSerLysLysLeuIleValAlaAspPheThAlaSerTrpCysGlyProCys 44  
 |||||  
 DB 99 GCTAATGAATCCAAACCTCTTGCGTGTGATTTCACGCGCTTCTTGCGTGTGACCATGT 158  
 QY 45 ArgPheMetAlaProValLeuAlaGluIleAlaLysLysThProGluLeuIlePheLeu 64  
 |||||  
 DB 159 GCTTTCATCGCTGCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 218  
 QY 65 LysValAspValAlaAspGluValArgProValAlaGluGluTrpSerIleGluAlaMetPro 84  
 |||||  
 DB 219 AAGGTTGATGACTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 278  
 QY 85 ThrPheLeuPheLeuLysAspGlyGluIleValAlaAspLysValAlaGlyAlaSerLysAsp 104  
 |||||  
 DB 279 ACCCTTCATGCTTTTGAAGGAGGAAGATTGACAAAGTTGAGACCAAGAAAGAT 338  
 QY 105 AspleuGlnAlaThrIleAlaLysHs 113  
 |||||  
 DB 339 GAGCTTCAGCTACCATTCACCAACAC 365

RESULT 8  
 LOCUS G73679 526 bp DNA linear STS 16-JUL-2002  
 DEFINITION R2488R etiolated leaf tissue of rice Oryza sativa STS genomic clone  
 R2488 sequencing direction=reverse, sequence tagged site.  
 ACCESSION G73679.1 GI:19697274  
 VERSION G73679.1 GI:19697274  
 KEYWORDS STS.  
 SOURCE Oryza sativa (indica cultivar-group)  
 ORGANISM Oryza sativa (indica cultivar-group)  
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;  
 Ehrhartoideae; Oryzaceae; Oryza.  
 1 (bases 1 to 526)  
 McCouch,S.R.  
 TITLE Oryza sativa STS  
 AUTHORS McCouch,S.R.  
 Journal Title Unpublished (2002)  
 COMMENT

Contact: Susan R. McCouch  
 Cornell University  
 Dept. of Plant Breeding, Ithaca, NY 14853-1901, USA  
 Tel: 6072550420  
 Fax: 6072556683  
 Email: srm4@cornell.edu  
 Primer A: M13 universal Forward GTAAACGACGGCCAGT  
 Primer B: M13 Universal Reverse AACACCTATGACCAAG  
 STS size: 526  
 Protocol:  
 Template: 20-100ng  
 Primer: 5pmol each  
 DNTPs: 40nmol  
 Tag polymerase: 5units  
 Total volume: 50ul  
 Buffer:  
 Tris-HCl: 100mM  
 KCl: 50mM  
 MgCl2: 15mM  
 Gelatin: 0.1%  
 PH: 8.3



[illegible][illegible]

Db 366 GACCTCAGAACACCATGTCGTAAGCAGCTGCGTCCACTGCTGATCTGCTTCC 422

RESULT 12

LOCUS NTRNA 698 bp mRNA linear PLN 15-FEB-1994

DEFINITION N.tabacum mRNA for thioredoxin.

ACCESSION X58527

VERSION X58527.1 GI:20046

KEYWORDS thioredoxin.

SOURCE Nicotiana tabacum (common tobacco)

ORGANISM Nicotiana tabacum

REFERENCE Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Asteridae; Lamiales; Solanales; Solanaceae; Nicotiana.

1 (bases 1 to 698)

REFERENCE Marty, I. and Meyer, Y.

LOCUS Nucleotide sequence of a cDNA encoding a tobacco thioredoxin

DB JOURNAL Plant Mol. Biol. 17 (1), 143-147 (1991)

MEDLINE 91329721

PUBMED 1868216

REFERENCE 2 (bases 1 to 698)

LOCUS Brugidou C., Marty, I., Chartier, Y. and Meyer, Y.

DEFINITION The Nicotiana tabacum genome encodes two cytoplasmic thioredoxin genes which are differentially expressed

ACCESSION Mol. Gen. Genet. 238 (1-2), 285-293 (1993)

93241165

8479434

REFERENCE 3 (bases 1 to 698)

LOCUS Meyer, Y.

DEFINITION Direct Submission

DB Submitted (14-FEB-1991) Y. Meyer, Laboratoire de Physiologie et Biologie, Moleculaire Vegetale, Universite av de Villeneuve, 66860 Perpignan, France

Gene product is probably cytoplasmic.

LOCATION/Qualifiers

1..698

/organism="Nicotiana tabacum"

/mol\_type="mRNA"

/strain="white Burley"

/db\_xref="taxon:4097"

/clone\_lib="plasmid PT219"

/dev\_stage="in vitro cells"

1..682

/evidence=experimental

79..459

/codon\_start=1

/product="thioredoxin"

/protein\_id="CAA1415.1"

/db\_xref="GI:20047"

/db\_xref="SWISS-PROT:P29449"

/translation="MAANDATSEGGVFGCHKEVNNERYKKGVETKLVVDFETAS WCGCRFIAPLADIARKMPHIVFLKVDVDELKVSASMSVEAMPTFVFINDKEVDR VVGAKREELQOTVIKHAAPATVTA"

468..473

polyA\_signal

511..517

polyA\_site

BASE COUNT 192 a 117 c 164 g 225 t

ORIGIN

Alignment Scores:

Pred. No.: 5.04e-37 Length: 698

Score: 404.00 Matches: 72

Percent Similarity: 83.78% Conservative: 21

Best Local Similarity: 64.86% Mismatches: 18

Query Match: 65.37% Indels: 0

DB: 8 Gaps: 0

US-09-786-715-8 (1-123) x NTRNA (1-698)

QY 5 GluGluGlyGlnValIleGlyValHisThrValAspGluTrpLysLeuGlnLeuGlnAsn 24

|||||

Db 106 GAGAGGAGACAAAGTTCGCTGCCACACAGGTTGAGGAATGGAACGAGTACTCTCAAGAA 165

QY 25 AlalysaspseryLysLeuIleValValAspPheThrAlaSerTrpCysGlyProCys 44

|||||

Db 166 GCGCTGAGACTAAGAAACGTGGTGGCTGATTTTACTGCTTATGGTGGCCCTTGC 225

QY 45 ArgPheMetAlaProValLeuAlaGluIleAlaLysLysTrpProGluLeuIlePheLeu 64

|||||

Db 226 CGTTTATTCGCCCAATCTTGTGACATTCGTAAGAAAGACGCCCATTTATATTCCTC 285

QY 65 LysValAspValAspGluValArgProValAlaGluIleLysTrpLeuAlaMetPro 84

|||||

Db 286 AAGGTGATGTTGATGAGAACTGATGACGCGGATGAGTGGAGGCAATGGCA 345

QY 85 ThrPheLeuPheLeuLysAspGlyValIleValAspLysValAlaGlyAlaSerLysAsp 104

|||||

Db 346 ACTTTTGTCTCATTAAGATGGAAGAAACATGGACACAGTGTGGTCCACAGAAAG 405

QY 105 AspleuGlnAlaThrIleAlaLysHisAlaSer 115

|||||

Db 406 GAGTTGCAGACACCATGATGTAAGCATCTGCT 438

AF483265

LOCUS AF483265 345 bp mRNA linear PLN 01-APR-2002

DEFINITION Populus tremula x Populus tremuloides thioredoxin H mRNA, complete cds.

ACCESSION AF483265

AF483265.1 GI:19851971

KEYWORDS

ORGANISM Populus tremula x Populus tremuloides

Populus tremula x Populus tremuloides

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids; eurosids I; Malpighiales; Salicaceae; Populus.

1 (bases 1 to 345)

REFERENCE Behm, M. and Jacquot, J. P.

LOCUS Isolation and characterization of thioredoxin h from poplar xylem

DB JOURNAL Plant Physiol. Biochem. 38 (5), 363-369 (2000)

2 (bases 1 to 345)

REFERENCE Behm, M. and Jacquot, J. P.

LOCUS Direct Submission

DB Submitted (13-FEB-2002) Faculte des Sciences, Universite Henri Poincare, Bd des Aiguillettes, Vandoeuvre 54506, France

LOCATION/Qualifiers

1..345

/organism="Populus tremula x Populus tremuloides"

/mol\_type="mRNA"

/db\_xref="taxon:47664"

1..345

/codon\_start=1

/product="thioredoxin H"

/protein\_id="NAL9994.1"

/db\_xref="GI:19851972"

/translation="MAEEGVYACHTVDYTKREHREKSGSOKLIVDFTSMPCKM IAPIFAEIAKKFPNVTFLKVDVDELKVAEBMVEAMPTFIFLKGKLVKTYGADKD GLPTLVAKHMTA"

BASE COUNT 101 a 61 c 95 g 88 t

ORIGIN

Alignment Scores:

Pred. No.: 4.78e-37 Length: 345

Score: 401.00 Matches: 73

Percent Similarity: 80.36% Conservative: 17

Best Local Similarity: 65.18% Mismatches: 22

Query Match: 64.89% Indels: 0

DB: 8 Gaps: 0

US-09-786-715-8 (1-123) x AF483265 (1-345)

QY 5 GluGluGlyGlnValIleGlyValHisThrValAspGluTrpLysLeuGlnLeuGlnAsn 24

|||||

Db 7 GAGAGGAGACAAAGTTCGCTGCCACACAGTGTGAGGAATGGAACGAGTACTCTCAAG 66

QY 25 AlalysaspseryLysLeuIleValValAspPheThrAlaSerTrpCysGlyProCys 44

```

Db      67 GGAAGAGGCTCAGAACTGATGTCGTGATTTACTGCTTCATGATGTCACCATGT 126
Oy      45 ArgpHemetaIaProValIeuaIaGluIleAlaIySlyThrProGluIlePheUeu 64
Db      127 AAATGATGTCACATCTTCGCGAGTGGGAGAAAGTTCCTCAATGTCATCTTGG 186
Oy      65 LysValaIspValaIspGluValaIspProValaIaGluIleUtyrSerIleGluIleMetPro 84
Db      187 AAGGTGATGTCGATGATGAATGAAGCTGTGCTGAGAGCTGGAATGTGGAGGATGCCA 246
Oy      85 ThrPheUeuPheUeuLysAspGlyGluIleValaIspLysValaIaGlyIleUtyrSerIleGlu 104
Db      247 ACTTTATTTCTTCCTGAAGATGCAAAATTAAGTGACAAACATGTGGCTGCTGATTAAGAT 306
Oy      105 AspleuGluIaIaThrIleAlaIySlySAlaSerAla 116
Db      307 GGCCTGCCAACACTGTTGCCAAAGCAGCAGCACTGCA 342

RESULT 14
TAE9762      596 bp      mRNA      linear      PLN 15-AUG-2001
LOCUS      Triticum aestivum mRNA for Chlorodoxin H.
ACCESSION      AJ009762.1 GI:4138593
VERSION      AJ009762.1 GI:4138593
KEYWORDS      Chlorodoxin H.
SOURCE      Triticum aestivum (bread wheat)
ORGANISM      Triticum aestivum
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
Poideae; Triticeae; Triticum.
REFERENCE      1
AUTHORS      Serrato,A.J., Crespo,J.L., Florencio,F.J. and Cejudo,F.J.
TITLE      Characterization of two chlorodoxins h with predominant
localization in the nucleus of aleurone and scutellum cells of
germinating wheat seeds
JOURNAL      Plant Mol. Biol. 46 (3), 361-371 (2001)
MEDLINE      21380673
PUBMED      11488482
REFERENCE      2 (bases 1 to 596)
AUTHORS      Cejudo,F.J.
TITLE      Direct Submission
JOURNAL      Submitted (24-JUL-1998) Cejudo F.J., Instituto de Bioquímica
Vegetal y Fotosíntesis, Universidad de Sevilla y CSIC, Avda Americo
Vespucio s/n, 41092-Sevilla, SPAIN
FEATURES
source      location/Qualifiers
1..596
/organism="Triticum aestivum"
/mol_type="mRNA"
/cultivar="Chinese Spring"
/db_xref="taxon:4565"
1..596
/gene="chlorodoxin H"
3'UTR      436..596
/gene="chlorodoxin H"
polyA_signal 462..466
/gene="chlorodoxin H"
BASE COUNT 161 a 127 c 182 g 126 t
ORIGIN
Alignment Scores:
Pred. No.: 2,64e-36 Length: 596
Score: 397.00 Matches: 68
Percent Similarity: 82.618 Conservative: 27
Best Local Similarity: 59.13% Mismatches: 20
Query Match: 64.24% Indels: 0
DB: 8 Gaps: 0
US-09-786-715-8 (1-123) x TAE9762 (1-596)
Oy      2 AlaguValaGluGluGlyValaIleGlyValaIsthrValaIspGluIurPlySleuGln 21
Db      85 GCGGGGGTGGGGGGGGAGGTGATCTCCCTCCACACCCCTGGAGAGTGCACATGCG 144

```

```

Oy      22 LeuGlnAsnAlaLysAspSerLysLysLeuIleValaIspPheThraIaSerTrpCys 41
Db      145 ATCGAGAGGCCAACCCGCCCAAGAGCTGGGGGTGATCTTCACTGCATCATGAGTGT 204
Oy      42 GlyProCysArgPheMetAlaProValIeuaIaGluIleAlaIySlyThrProGluIleu 61
Db      205 GGACCATGGCGCATATGCTCCAGTTTCGGCTGATCTCGCCACAGAAATGCCAATGCT 264
Oy      62 IlePheUeuLysValaIspValaIspGluValaIspProValaIaGluIleUtyrSerIleGlu 81
Db      265 GTTTTCCTCAAGTGCATGTCATGATGAAGACCCATTCGAGAGCAATTCAGCGTTGAG 324
Oy      82 AlameIProThrPheUeuPheUeuLysAspGlyGluIleValaIspLysValaIaGlyIle 101
Db      325 GCCATGCCAACCTCTCTGTCATTTAAGAGAGAGATGTCAGAGCAGGGGTGGAGCT 384
Oy      102 SerLysAspAspleuGluIaIaThrIleAlaIySlySAlaSerAla 116
Db      385 ATCAAGAGGAACTGACGACAAAGTGGGCTACACGGCGGCC 429

RESULT 15
AY271308      784 bp      mRNA      linear      PLN 12-MAY-2003
LOCUS      Citrus x paradisi chlorodoxin H mRNA, complete cds.
DEFINITION      Citrus x paradisi chlorodoxin H mRNA, complete cds.
ACCESSION      AY271308
VERSION      AY271308.1 GI:30575685
KEYWORDS
SOURCE      Citrus x paradisi
ORGANISM      Citrus x paradisi
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
Rosids; eustosids II; Sapindales; Rutaceae; Citrus.
REFERENCE      1 (bases 1 to 784)
AUTHORS      Porat,R., Ben-Yephet,M. and Drobny,S.
TITLE      Isolation of a chlorodoxin H cDNA from grapefruit induced by yeast
and fungal treatments
JOURNAL      unpublished
REFERENCE      2 (bases 1 to 784)
AUTHORS      Porat,R., Ben-Yephet,M. and Drobny,S.
TITLE      Direct Submission
JOURNAL      Submitted (08-APR-2003) Postharvest Science, ARO, The Volcani
Center, Derech Hamacabim, Bet Dagan 50250, Israel
FEATURES
source      location/Qualifiers
1..784
/organism="Citrus x paradisi"
/mol_type="mRNA"
/db_xref="taxon:37656"
190..561
/codon_start=1
/product="chlorodoxin H"
/protein_id="AAP33009.1"
/db_xref="GI:30575686"
/translation="MAEEGVISCHTVESNMEOLOGIAAKLIIVDFPASCPCKL
MSPILSELAKLPVAVIFLKVDYDELQSAEVAEVAEMPFVLTGKYLRIYAKKD
EQQLAVEKHATIVENATPANA"
BASE COUNT 236 a 141 c 186 g 214 t 7 others
ORIGIN
Alignment Scores:
Pred. No.: 1,05e-35 Length: 784
Score: 393.00 Matches: 70
Percent Similarity: 80.678 Conservative: 26
Best Local Similarity: 58.82% Mismatches: 23
Query Match: 63.59% Indels: 0
DB: 8 Gaps: 0
US-09-786-715-8 (1-123) x AY271308 (1-784)
Oy      5 GluGluGlyGlnValaIleGlyValaIsthrValaIspGluIurPlySleuGlnLeuGlnAsn 24
Db      196 GAGGAGGGGCAAGTAGTATGACGTGCCACACTGTGATCTCGAACAGCAGACAGCTCAAAAG 255

```

```

QY      25 AlalysaspSerLysLysLeuIleValValaspPheThrAlaSerTrpCysGlyProCys 44
      256 GGCATTGGCGGAAGAACTGATAGTGTGATTTTACGGCTTCATGTGCCCCCATGC 315
QY      45 ArgPheMetAlaProValLeuAlaGluIleAlaLysLysThrProGluLeuIlePheLeu 64
      316 AAGTTGATGTCTCCAAATTGTGAGTGAATTGGCGAAGAGCTCCCGCTGCATATTTCTTG 375
Db
QY      65 LysValaspValaspGluValArgProValAlaGluGluIleGluIleGluAlaMetPro 84
      376 AAGGTGATGTGTGATGATTCAGTCCGCTCCGCTGAGGATGGCTGTGGAGCCATGCG 435
QY      85 ThrPheLeuPheLeuLysaspGlyGluIleValaspLysValAlaGlyAlaSerLysasp 104
      436 ACCTTTGTACTGACGAAAGAGCGCAAGTCTCTGAGAGGATTTGGGAGCGCAAGAAAGAT 495
Db
QY      105 AspleuGluAlaThrIleAlaLysHisAlaSerAlaValAlaAlaAlaSerSerSer 123
      496 GAGCTGCAGCTGGCTGTGCAAAAGCACGCAACACTGTGGAATGCAACTACTGCT 552

```

Search completed: August 17, 2003, 21:41:58  
 Job time : 1904.81 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - nucleic search, using frame\_plus\_p2n model

Run on: August 17, 2003, 17:46:53 ; Search time 172.913 Seconds

(without alignments)  
1920.219 Million cell updates/sec

Title: US-09-786-715-8

Perfect score: 618

Sequence: 1 MAEEVEGQIVGVHTEDEMKL.....DDLQATIAKASVAASS 123

Scoring table:

BLOSUM62  
Xgapop 10.0 , Xgapext 0.5  
Ygapop 10.0 , Ygapext 0.5  
Fgapop 6.0 , Fgapext 7.0  
Delop 6.0 , Delext 7.0

Searched: 2552756 seqs, 1349719017 residues

Total number of hits satisfying chosen parameters: 5105512

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:  
-MODE=frame-p2n.model -DEV=rlp  
-Q=/cgn2.1/USPTO.spool/US09786715/runat.11082003.150513.6025/app\_query.fasta.1.1052  
-DB=N\_Geneseq.19Jun03 -QFMT=fastap -SUFFIX=ring -MINMATCH=0.1 -LOOPEL=0  
-LOOPEXT=0 -UNITS=bits -START=1 -END=-1 -MATRIX=biosum62 -TRANS=human40.cdi  
-LIST=45 -DOCALLGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15  
-MODE=LOCAL -OUTFMT=ptc -NORM=ext HEAPSIZ=500 -MINLEN=0 -MAXLEN=200000000  
-USER=US09786715.ecgn.1.1.874.6runat.11082003.150513.6025 -NCP=6 -ICPU=3  
-NO\_MAP -LAKEOUTERY -NEG\_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG  
-DEV\_TIMEROUT=120 -WARN\_TIMEROUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6  
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

N\_Geneseq.19Jun03.\*  
1: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1980.DAT.\*  
2: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1981.DAT.\*  
3: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1983.DAT.\*  
4: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1983.DAT.\*  
5: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1984.DAT.\*  
6: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1985.DAT.\*  
7: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1986.DAT.\*  
8: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1987.DAT.\*  
9: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1988.DAT.\*  
10: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1989.DAT.\*  
11: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1990.DAT.\*  
12: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1991.DAT.\*  
13: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1992.DAT.\*  
14: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1993.DAT.\*  
15: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1994.DAT.\*  
16: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1995.DAT.\*  
17: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1996.DAT.\*  
18: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1997.DAT.\*  
19: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1998.DAT.\*  
20: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1999.DAT.\*  
21: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA2000.DAT.\*  
22: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT.\*  
23: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA2001B.DAT.\*  
24: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA2002.DAT.\*  
25: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA2003.DAT.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed,

and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Match	Length	DB	ID	Description
1	618	100.0	601	21	AAZ51740	Glycine max thior
2	565	91.4	390	25	ABX19403	Human GDP-mannose
3	515	83.3	392	25	ABX18045	Human GDP-mannose
4	446	72.2	574	21	AAZ51738	Catalpa speciosa t
5	425	68.8	402	25	ABX21664	Human GDP-mannose
6	418	67.6	738	21	AAZ51739	Glycine max thior
7	413	66.8	270	25	ABX31386	Human GDP-mannose
8	411	66.5	3888	24	ABX53097	DNA encoding Thior
9	411	66.5	3888	24	ABN89581	Phaseolin promoter
10	411	66.5	4935	24	ABN89586	Promoter-Oleosin t
11	410	66.3	345	24	ABZ12358	Arabidopsis thalia
12	410	66.3	345	24	ABN89587	Arabidopsis thalia
13	410	66.3	470	21	AAZ51781	Arabidopsis thalia
14	410	66.3	561	21	AAZ51781	Arabidopsis thalia
15	410	66.3	563	21	AAZ51781	Arabidopsis thalia
16	410	66.3	3129	24	ABX53095	DNA encoding thior
17	410	66.3	3129	24	ABN89579	Phaseolin promoter
18	410	66.3	3888	24	ABX53096	DNA encoding Oleos
19	410	66.3	3888	24	ABN89580	Phaseolin promoter
20	406	65.7	686	15	AAZ78205	Gene coding for pr
21	406	65.7	687	21	AAZ66375	Rice thiorodoxin h
22	404.5	65.5	509	22	AAH87768	Peppermint plant o
23	404	65.4	614	21	AAZ51741	Vernonia mespilifo
24	392	63.4	393	17	AAZ10451	Hard wheat thior
25	392	63.4	393	21	AAZ62457	Wheat thiorodoxin
26	392	63.4	653	16	AAZ09783	Plant SAR gene pl.
27	392	63.4	653	20	AAV62799	Tobacco SAR CHX in
28	392	63.4	653	20	AAV81683	Nucleotide sequenc
29	389	62.9	369	21	AAZ61537	Wheat thiorodoxin
30	388	62.8	382	21	AAZ62456	Soft wheat thior
31	387	62.8	384	17	AAZ10450	Plant microsatelli
32	387	62.6	419	21	AAZ13142	Arabidopsis thalia
33	380.5	61.6	360	24	ABZ12359	Arabidopsis thalia
34	376	60.8	370	21	AAZ62455	Barley thiorodoxin
35	376	60.8	560	21	AAZ62455	Arabidopsis thalia
36	374.5	60.6	590	21	AAZ38792	Arabidopsis thalia
37	374	60.5	357	24	AAZ13931	Arabidopsis thalia
38	374	60.5	480	21	AAZ36542	Arabidopsis thalia
39	374	60.5	563	21	AAZ34211	Arabidopsis thalia
40	374	60.5	563	21	AAZ34211	Arabidopsis thalia
41	373	60.4	524	21	AAZ33829	Arabidopsis thalia
42	367	59.4	576	25	ABX56868	Arabidopsis thalia
43	361	58.9	260	25	ABX30882	Human GDP-mannose
44	361	58.4	267	25	ABX31078	Human GDP-mannose
45	360	58.3	353	21	AAZ31118	Plant microsatelli

## ALIGNMENTS

RESULT 1	AAZ51740	standard; cDNA: 601 BP.
ID	AAZ51740;	
AC	AAZ51740;	
XX		
DT	04-JUL-2000	(first entry)
XX		
DE	Glycine max thiorodoxin	cdna-2.
XX		
KW	Glycine max thiorodoxin; clone sfil.pK0029.e2; chimeric gene; soybean;	
KM	transgenic plant; seed storage protein; allergenicity; ss.	
XX		
OS	Glycine max.	
XX		
FT	Key	Location/Qualifiers
CDS		37..408





US-09-786-715-8 (1-123) x ABX19403 (1-390)

QY 1 MetAlaGluValGluGluLysValIleGlyValHisThrValAspGluTrpLysLeu 20  
DB 57 ATGGCGAAGTGAAGAGGACAGTCATCGGCGTCCACACCTTGATGATGGAGCTG 116  
QY 21 GlnLeuGlnAsnAlaLysAspSerLysLysLeuIleValValAspPheThrAlaSerTyr 40  
DB 117 CAAGTCAGAAATGCAAAAGACTCCAAAACATGATGTTGTGTGATTTTACTGCTTCTG 176  
QY 41 CysGlyProCysArgPheMetAlaProValLeuAlaGluIleAlaLysIleThrProGlu 60  
DB 177 TGGGTCCATGGCGCTTTATGCGCCAGTTCTTGAGAGATGCAAAAGAACTCCTGAA 236  
QY 61 LeuIlePheLeuLysValAspValAspGluValArgProValAlaGluGluTyrSerTyr 80  
DB 237 TTGATCTTCCTCAAGTGAATGATGATGAGGCGCTGCTAGGAATATTCCTATT 296  
QY 81 GluAlaMetProThrPheLeuPheLeuLysAspGlyGluIleValAspLysValAlaGly 100  
DB 297 GAGGCCATGCAACCTTCTCTTCTTGAAGAATGGCGCAATCGTGACAAAGTGTGGT 356  
QY 101 AlaSerLysAspAspLeuGlnAlaThrIleAla 111  
DB 357 GCTAGTAAGATGACCTTCAGCCACCATAGCC 389

RESULT 3  
ABX18045  
ID ABX18045 standard; cDNA; 392 BP.  
XX  
XX ABX18045;  
DT 10-FEB-2003 (first entry)  
XX  
DE Human GDP-mannose 4,6-dehydratase (GM4,6D) DNA #102.  
XX  
XX Human: GDP-mannose 4,6-dehydratase; GM4,6D; gene; ss; inflammation;  
KM cellular fucosylation; glycoconjugate fucosylation; transplant rejection;  
KM arthritis; asthma; sepsis; reperfusion injury; stroke; infection;  
KM complex carbohydrate; gene replacement therapy; immunosuppressive;  
KM antiinflammatory; antiarthritic; antibacterial; cerebroprotective;  
KM antiasthmatic; vasotropic.  
XX  
OS Homo sapiens.  
XX  
PN US2002110548-A1.  
XX  
PD 15-AUG-2002.  
XX  
PF 11-JUN-2001; 2001US-0878574.  
XX  
PR 22-NOV-1996; 96US-0753233.  
PR 03-DEC-1997; 97US-0984246.  
PR 09-SEP-1998; 98US-0149674.  
PR 14-JUN-1999; 99US-0333177.  
XX  
XX (GENW ) GENETICS INST INC.  
XX  
PI Sullivan F, Kriz R, Kumar R;  
XX  
XX WPI; 2003-066673/06.  
XX  
PT New composition comprising GDP-mannose 4,6-dehydratase (GM4,6D)  
PT peptide, for manufacturing complex carbohydrates, or as targets for  
PT screening GM4,6D antagonists for treating e.g. arthritis, or transplant  
PT rejection  
XX  
PS Disclosure: SEQ ID NO 104; 6pp; English.  
XX  
XX  
CC The invention relates to a composition comprising a human GDP-mannose  
CC 4,6-dehydratase (GM4,6D) peptide. The peptide is useful for identifying  
CC GM4,6D inhibitors. GM4,6D inhibitors are useful for reducing inflammation  
CC in a mammalian subject and for treating or ameliorating diseases affected

CC by the level of cellular fucosylation or diseases affected by the  
CC fucosylation of glycoconjugates. These diseases include arthritis,  
CC transplant rejection, asthma, sepsis, reperfusion injury, stroke or  
CC infection. The GM4,6D peptide or a polynucleotide encoding it is also  
CC useful for manufacturing complex carbohydrates and as targets for  
CC screening small molecule antagonists of the activity of the enzyme. The  
CC polynucleotide is useful in developing an assay for defects in the  
CC enzyme, as well as in gene replacement therapy. Sequences  
CC ABX1942-ABX1944 and ABX1947-ABX33716 represent DNA molecules encoding  
CC human GM4,6D peptides of the invention.  
CC Note: The sequence data for this patent did not form part of the printed  
CC specification but was obtained in electronic format directly from USPTO  
CC at [seqdata.uspto.gov/sequence.html](http://seqdata.uspto.gov/sequence.html).  
XX

SEQ Sequence 392 BP; 99 A; 80 C; 107 G; 106 T; 0 other;

Alignment Scores:

Pred. No.:	Length:	392
Score:	515.00	98
Percent Similarity:	92.31%	10
Best Local Similarity:	83.76%	9
Query Match:	83.33%	0
DB:	25	Gaps:

US-09-786-715-8 (1-123) x ABX18045 (1-392)

QY 5 GluGluGlyGlnValIleGlyValHisThrValAspGluTrpLysLeuGlnAsn 24  
DB 12 GAAGAGGTCAGAGTGTGCGCGTCCACACCGTGTGATGCGGAAACGACAGTGCAGAAAT 71  
QY 25 AlaLysAspSerLysLysLeuIleValAspPheThrAlaSerTyrCysGlyProCys 44  
DB 72 GCAAAAGACTCCAAAATTTGATTGGTGATTTACTGCTCGGTGGTGGTCCATGC 131  
QY 45 ArgPheMetAlaProValLeuAlaGluIleAlaLysIleThrProGluLeuIlePheLeu 64  
DB 132 CGTTTATATGCCCGCACTGTCGAGAGATGCTAGCAACATCCCAAGATCTTCTC 191  
QY 65 LysValAspValAspLysValArgProValAlaGluGluTyrSerIleGluAlaMetPro 84  
DB 192 AAGGTGATGTGATGAAGTGAAGCGCTGTCTGAGGAATATTCCATTGAGGCCATGCCA 251  
QY 85 ThrPheLeuPheLeuLysAspGlyGluIleValAspLysValValGlyLaserLysAsp 104  
DB 252 ACCTTCCTCTTCTTGAAGATGCAAGATCGTGAATAGCTGTGTGCTAAGAAAGAG 311  
QY 105 AspLeuGlnAlaThrIleAlaLysHisAlaSerAlaValAlaAlaLaser 121  
DB 312 GAGCTCAGACTCACCATAGCCAAAGCATGTATCTGCTGCTGCTCTTCT 362

RESULT 4  
AAZ51738  
ID AAZ51738 standard; cDNA; 574 BP.  
XX  
XX AAZ51738;  
XX  
DT 04-JUL-2000 (first entry)  
XX  
DE Catalpa speciosa thiorodoxin cDNA.  
XX  
DE Catalpa speciosa thiorodoxin; clone ncs.pk0010.e3; chimeric gene;  
KM transgenic plant; seed storage protein; allergenicity; ss.  
XX  
OS Catalpa speciosa.  
XX  
XX  
XX Key Location/Qualifiers  
FH 63..419  
FT CDS /tag= a  
FT /product= "thiorodoxin"  
XX  
PN WO200014239-A2.  
XX  
PD 16-MAR-2000.

XX 07-SEP-1999; 99WO-US20420.  
PF 08-SEP-1998; 98US-0099501.  
PR (DUPO ) DU PONT DE NEMOURS & CO E. I.  
XX Allen SM, Thorpe CJ, Lu AL;  
XX WPI: 2000-256987/22.  
DR P-PSDB: AAY70480.  
XX  
XX New isolated polynucleotide encoding thioredoxin polypeptide is useful  
PT for producing transgenic plants with an altered level of thioredoxin -  
XX  
XX Claim 3; Page 29; 33pp; English.  
XX  
XX The present cDNA sequence encodes Catalpa speciosa thioredoxin  
CC protein. The cDNA was derived from clone ncs.pk0010.63, which was  
CC isolated from a cDNA library prepared from C. speciosa developing seed  
CC tissue. Chimeric genes encoding all or a portion of the thioredoxin  
CC protein, in sense or antisense orientation are constructed, wherein  
CC expression of the chimeric gene results in production of altered levels  
CC of the thioredoxin protein in a transformed host cell. Thioredoxin is  
CC involved in the disassembly of seed storage proteins during germination  
CC by reducing S-S bonds and in the bread making process. Over expression of  
CC thioredoxin in cereals may reduce the allergenicity of any transgenic  
CC protein engineered into cereal crops with high sulfhydryl content.  
XX  
XX Sequence 574 BP; 171 A; 91 C; 157 G; 155 T; 0 other;  
SO  
Alignment Scores:  
Pred. No.: 7.76e-51 Length: 574  
Score: 446.00 Matches: 80  
Percent Similarity: 85.59% Conservative: 21  
Best Local Similarity: 67.80% Mismatches: 17  
Query Match: 72.17% Indels: 0  
DB: Gaps: 0  
US-09-786-715-8 (1-123) x AA251738 (1-574)  
QY 1 MetAlaGluValAlaGluGluGluValAlaIleGlyValHisThrValAspGluTrpLysLeu 20  
DB 63 ATGGCTCTTCAGAAAGGAGGCAAGTGTGCGTCCACCTCCGACGACTGGAAGAG 122  
QY 21 GluLeuGlnAsnAlaLysAspSerLysLysLeuIleValValAspPheThrAlaSerTrp 40  
DB 123 CAGTCCAGAAAGGCTGTGACTCTAAGAAACTGGTGTGAATAGACTTCACGGCTTCCG 182  
QY 41 CysGlyProCysArpPheMetAlaProValLeuAlaGluIleAlaLysLysThrProGlu 60  
DB 183 TGCAGACCAATGCCCTTCATGTCTCAATCTTGGCTGAGAGGCCAAGAACACCCCAT 242  
QY 61 LeuIlePheLeuLysValAspValAspGluValArgProValAlaGluGluTrpSerIle 80  
DB 243 GTCAATATTCCTGAAGAGTGCAGGTGATGAACCTCAAGACTGTGCTGAGAAATTCAAA 302  
QY 81 GluAlaMetProThrPheLeuPheLeuLysAspGlyGluIleValAlaAspLysValAlaGly 100  
DB 303 GAGGCTAATGCCGACCTTCGTCTTCTCAAGGAAGGAAGAAGGCAAGAGCTTGTGGA 362  
QY 101 AlaSerLysAspAspLeuGlnAlaThrIleAlaLysHisAlaSerAlaValAla 118  
DB 363 GCAAGGAAGGAGGAATTGCAAGGCCACAGTTGAGAAACATGCGCTATCAGCTGCT 416  
RESULT 5  
ABX21664  
ID ABX21664 standard; cDNA; 402 BP.  
XX  
XX ABX21664;  
XX 10-FEB-2003 (first entry)

DE Human GDP-mannose 4,6-dehydratase (GM4,6D) DNA #3721.  
XX  
XX Human; GDP-mannose 4,6-dehydratase; GM4,6D; gene; ss; inflammation;  
KW cellular fucosylation; glycoconjugate fucosylation; transplant rejection;  
KW arthritis; asthma; sepsis; reperfusion injury; stroke; infection;  
KW complex carbohydrate; gene replacement therapy; immunosuppressive;  
KW antiinflammatory; antiarthritic; antibacterial; cerebroprotective;  
KW antiasthmatic; vasotropic.  
XX  
XX Homo sapiens.  
XX  
XX US2002110548-A1.  
XX  
XX 15-AUG-2002.  
XX  
XX 11-JUN-2001; 2001US-0878574.  
XX  
XX 22-NOV-1996; 96US-0753233.  
XX 03-DEC-1997; 97US-0984246.  
XX 09-SEP-1998; 98US-0149674.  
XX 14-JUN-1999; 99US-0333177.  
XX  
XX (GENV ) GENETICS INST INC.  
XX  
XX Sulliyvan F, Kriz R, Kumar R;  
XX  
XX WPI: 2003-066673/06.  
XX  
XX New composition comprising GDP-mannose 4,6-dehydratase (GM4,6D)  
PT peptide, for manufacturing complex carbohydrates, or as targets for  
PT screening GM4,6D antagonists for treating e.g. arthritis, or transplant  
PT rejection  
XX  
XX Disclosure; SEQ ID NO 3723; 6pp; English.  
XX  
XX The invention relates to a composition comprising a human GDP-mannose  
CC 4,6-dehydratase (GM4,6D) peptide. The peptide is useful for identifying  
CC GM4,6D inhibitors. GM4,6D inhibitors are useful for reducing inflammation  
CC in a mammalian subject and for treating or ameliorating diseases affected  
CC by the level of cellular fucosylation or diseases affected by the  
CC fucosylation of glycoconjugates. These diseases include arthritis,  
CC transplant rejection, asthma, sepsis, reperfusion injury, stroke or  
CC infection. The GM4,6D peptide or a polynucleotide encoding it is also  
CC useful for manufacturing complex carbohydrates and as targets for  
CC screening small molecule antagonists of the activity of the enzyme. The  
CC polynucleotide is useful in developing an assay for defects in the  
CC enzyme, as well as in gene replacement therapy. Sequences  
CC ABX17942-ABX17944 and ABX17947-ABX3716 represent DNA molecules encoding  
CC human GM4,6D peptides of the invention.  
CC Note: The sequence data for this patent did not form part of the printed  
CC specification but was obtained in electronic format directly from USPTO  
CC at [seqdata.uspto.gov/sequence.html](http://seqdata.uspto.gov/sequence.html).  
XX  
XX Sequence 402 BP; 113 A; 76 C; 122 G; 91 T; 0 other;  
SO  
Alignment Scores:  
Pred. No.: 3.31e-48 Length: 402  
Score: 425.00 Matches: 77  
Percent Similarity: 86.49% Conservative: 19  
Best Local Similarity: 69.37% Mismatches: 15  
Query Match: 68.77% Indels: 0  
DB: Gaps: 0  
US-09-786-715-8 (1-123) x ABX21664 (1-402)  
QY 5 GluGluGluGluValAlaIleGlyValHisThrValAspGluTrpLysLeuGlnLeuGlnAsn 24  
DB 49 GAAGAGGACAAAGTGTGCTGCACACCTTAAGAGAGTGAAGAACATCTCAAGAG 108  
QY 25 AlaLysAspSerLysLysLeuIleValValAspPheThrAlaSerTrpCysGlyProCys 44  
DB 109 GGAGAAAGTCCCAAGAAATTCATGTGTGATTTTACTGCTTGTGTGGCGTCCATGC 168

QY	45	ArgPheMetAlaProValLeuAlaGluIleAlaLysLysThrProIleuLeuIlePheLeu	64
Db	169	CGTTTCATTGGCCCCAATTCTTGACAGATTGTGCTAAAAAAGTTGGCAATGTCACTTCCTC	228
QY	65	LysValAspValAspGluValArgProValAlaGluGluTyrSerIleGluAlaMetPro	84
Db	229	AACGTGGATGTGGATGAATTTGGAGACTGTCTTCCACAGCAATGGGGAATTGAGCTATGGCCA	288
QY	85	ThrPheLeuPheLeuLysAspGlyGluIleValAspLysValAlaGlyAlaSerLysAsp	104
Db	289	ACCTTCCTGTCTTGAAAGAACGCAAGCTGTGGACACAGGTTGTGGTGCCACAAAGAC	348
QY	105	AspLeuGlnAlaThrIleAlaLysHisAlaSer	115
Db	349	GAGCTGCAATTGACCCCTACGACACATGGCCGA	381
RESULT 6			
ID	AA251739	standard: cDNA; 738 BP.	
AC	AA251739;		
DT	04-JUL-2000	(first entry)	
DE		Glycine max thioredoxin cDNA-1.	
XX			
XX		Glycine max thioredoxin; clone sahlc.pk001.117; chimeric gene; soybean;	
XX		transgenic plant; seed storage protein; allergenicity; ss.	
OS		Glycine max.	
XX			
EH	Key	Location/Qualifiers	
FT	CDS	68..430	
FT		/*tag= "a	
FT		/product= "Thioredoxin"	
XX			
PM	WO200041239-A2.		
XX			
PD	16-MAR-2000.		
XX			
XX	07-SEP-1999;	99WO-US20420.	
PF			
XX	08-SEP-1998;	98US-0099501.	
PR			
XX			
PA	(DUPO ) DU PONT DE NEMOURS & CO E I.		
XX			
PI	Allen SM, Thorpe CJ, Lu AL;		
DR	WPI: 2000-256987/22.		
DR	P-PSDB; AAY70481.		
XX			
PT	New isolated polynucleotide encoding thioredoxin polypeptide is useful		
PT	for producing transgenic plants with an altered level of thioredoxin -		
XX			
XX	Claim 3; Page 30; 33pp; English.		
XX			
CC	The present cDNA sequence encodes Glycine max thioredoxin protein. The		
CC	cDNA was derived from clone sahlc.pk001.117, which was isolated from a		
CC	cDNA library prepared from soybean tissue sprayed with authority		
CC	herbicide. Chimeric genes encoding all or a portion of the thioredoxin		
CC	protein, in sense or antisense orientation are constructed, wherein		
CC	expression of the chimeric gene results in production of altered levels		
CC	of the thioredoxin protein in a transformed host cell. Thioredoxin is		
CC	involved in the disassembly of seed storage proteins during germination		
CC	by reducing S-S bonds and in the bread making process. Over expression of		
CC	thioredoxin in cereals may reduce the allergenicity of any transgenic		
CC	protein engineered into cereal crops with high sulphydryl content.		
XX			
SQ	Sequence 738 BP; 260 A; 120 C; 146 G; 212 T; 0 other;		
Alignment Scores:	6.89e-47	Length: 738	
Pred. No.:	418.00	Matches: 78	
Score:			

Percent Similarity:	83.19%	Conservative:	21
Best Local Similarity:	65.55%	Mismatches:	20
Query Match:	67.64%	Indels:	0
DB:	21	Gaps:	0

US-09-786-715-8 (1-123) x AA251739 (1-738)

QY	5	GlucIuGlYglYlnValIleGlYValHlStrhValAspGlUtrPrLysleuGlInleuGlnAsn	24
DB	83	GAAGAGGGAGCAAGTCATTAGCTGCACACCGTTGAAGAAATGAGACGATCACTCCAGAG	147
QY	25	AlAlaYasPserLysLysLeuIleValValAspPheThrAlaSerTTPcysGlyProCys	44
DB	143	GGCAACCAACCAAGAAACATCATGTGGATTACTGCTTCTTGGGTGGAGACCATGC	202
QY	45	ArgPheMetIleProValLeuAlaGluIleAlaLysLysThrProGluIlePheLeu	64
DB	203	CGTTTCTTCTACCATTCATTCTTGGCTGACGTGGCTTAAGAAAGTTACAAAGTGTCTATTCTTA	265
QY	65	LysValAspValAspGluValArgProValAlaGluGluTrpSerIleGluAlaMetPro	84
DB	263	AAGGTGATGTGGACGAATTAAGAGTGTTCTCAAGATTTGGCTATTGAGGCTATGCC	322
QY	85	ThrPheLeuPheLeuLysAspGlyIuIleValAspLysValValGlyAlaSerLysAsp	104
DB	323	ACTTTTGTTGTGTGAAGAGGGACGCTTGGACAAAGTGTGGAGCAAGAGCAT	382
QY	105	AspLeuGlnAlaThrIleAlaLysHlAlaSerAlaValAlaAlaAlaSerSerSer	123
DB	383	GAGCTGCACGACAAATACAGAACATGTGGCTTCAGCTAGTCTTAATCTAGTCA	439

RESULT 7  
ABX31386 standard; cDNA; 270 BP.

ID	ABX31386	standard; cDNA; 270 BP.
AC	ABX31386;	
XX		
DT	11-FEB-2003	(first entry)
XX		
DE	Human GDP-mannose 4,6-dehydratase (GM4,6D) DNA #13443.	
XX		
KW	Human; GDP-mannose 4,6-dehydratase; GM4,6D; gene; ss; inflammation;	
KW	cellular fucosylation; glycoconjugate fucosylation; transplant rejection;	
KW	arthritis; asthma; sepsis; reperfusion injury; stroke; infection;	
KW	complex carbohydrate; gene replacement therapy; immunosuppressive;	
KW	antiinflammatory; antiarthritic; antibacterial; cerebroprotective;	
KW	antiaesthetic; vasotropic.	
XX		
OS	Homo sapiens.	
PN	US2002110548-A1.	
XX		
PD	15-AUG-2002.	
XX		
PE	11-JUN-2001; 2001US-0878574.	
XX		
PR	22-NOV-1996; 96US-0753233.	
PR	03-DEC-1997; 97US-0984246.	
PR	09-SEP-1998; 98US-0149674.	
XX	14-JUN-1999; 99US-0333177.	
XX		
PA	(GENM ) GENETICS INST INC.	
PI	Sullivan F, Kriz R, Kumar R;	
DR	WPI: 2003-066673/06.	
XX		
PT	New composition comprising GDP-mannose 4,6-dehydratase (GM4,6D)	
PT	peptide, for manufacturing complex carbohydrates, or as targets for	
PT	screening GM4,6D antagonists for treating e.g. arthritis, or transplant	
PT	rejection	
XX		
XX	Disclosure; SEQ ID NO 13445; 69p; English.	

XX The invention relates to a composition comprising a human GDP-mannose  
CC 4,6-dehydratase (GM4,6D) peptide. The peptide is useful for identifying  
CC GM4,6D inhibitors. GM4,6D inhibitors are useful for reducing inflammation  
CC in a mammalian subject and for treating or ameliorating diseases affected  
CC by the level of cellular fucosylation or diseases affected by the  
CC fucosylation of glycoconjugates. These diseases include arthritis,  
CC transplant rejection, asthma, sepsis, reperfusion injury, stroke or  
CC infection. The GM4,6D peptide or a polynucleotide encoding it is also  
CC useful for manufacturing complex carbohydrates and as targets for  
CC screening small molecule antagonists of the activity of the enzyme. The  
CC polynucleotide is useful in developing an assay for defects in the  
CC enzyme, as well as in gene replacement therapy. Sequences  
CC ABX17942-ABX17944 and ABX17947-ABX3716 represent DNA molecules encoding  
CC human GM4,6D peptides of the invention.  
CC Note: The sequence data for this patent did not form part of the printed  
CC specification but was obtained in electronic format directly from USPTO  
CC at [seqdata.uspto.gov/sequence.html](http://seqdata.uspto.gov/sequence.html).

XX Sequence 270 BP; 69 A; 53 C; 71 G; 77 T; 0 other;

Alignment Scores:

Pred. No.: 8,03e-47 Length: 270  
Score: 413.00 Matches: 80  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 66.83% Indels: 0  
DB: 25 Gaps: 0

US-09-786-715-8 (1-123) x ABX31386 (1-270)

QY 1 MetAlaGlValAlGluGluGluValAlIleGlyValAlHisThrValAspGluTrpLysLeu 20  
DB 29 ATGGCTGAAGTGAAGGAGGAGGAGTATCGCGCCACACCGTTGATGAGTGAAGCTG 88  
QY 21 GluLeuGlnAsnAlaLysAspSerLysLysLeuIleValAlAspPheThrAlaSerTrp 40  
DB 89 CAACTCCAGATGCAAAAGACTCCAAAAACTGATTTGGTGGATTTTACTGCTCCCTGG 148  
QY 41 CysGlyProCysATGCGPheMetAlaProValLeuAlaGluIleAlaLysLysThrProGlu 60  
DB 149 TGTGCTCATGCTCCGTTTATGCCCCACTTCTTGCAAGATTCGAAGAAACCTCTGAA 208  
QY 61 LeuIlePheLeuLysValAspValAspGluValAlArgProValAlaGluGluTrpSerIle 80  
DB 209 TTGATCTTCTCAAGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 268

RESULT 8

ABSS3097  
ID ABSS3097 standard; DNA; 3888 BP.

XX AC ABSS3097;

DT 29-NOV-2002 (first entry)

XX DNA encoding Thioresdoxin-oleosin fusion protein.

KW Thioresdoxin; thioresdoxin reductase; gene expression; oleosin;  
KW oil body; oleosin-thioresdoxin fusion protein; gene; ds.

XX Arabidopsis thaliana.

OS Brassica napus.

OS Synthetic.

XX Key location/Qualifiers

FT CDS 1555..2658  
FT /tag= a  
FT /product= "Thioresdoxin-oleosin fusion protein"

FT exon 1555..2250  
FT /tag= b  
FT /number= 1

FT Intron 2251..2489  
FT /tag= c

FT /number= 1  
FT exon 2490..2658  
FT /tag= d  
FT /number= 2

US2002088025-A1.

04-JUL-2002.

03-JUL-2001; 2001US-0897425.

22-FEB-1991; 91US-0659835.

16-NOV-1993; 93US-0142418.

30-DEC-1994; 94US-0366783.

25-APR-1997; 97US-0846021.

18-DEC-1998; 98US-0210843.

(MOLO/) MOLONEY M. M.  
(DALM/) DALMIA B. K.

MoLoney NM, Dalmlia BK;

WPI: 2002-635723/68.

Expressing protein, by introducing chimeric nucleotide regulatory  
sequence, sequence encoding fusion protein, having sequence encoding  
protein, oleosin gene and sequence encoding termination region and  
producing protein

Example 21: Fig 14; 69pp; English.

XX The invention describes a method of expressing thioresdoxin or thioresdoxin  
CC reductase (I) in the oil body of a host cell using an oil body protein  
CC gene. The method involves introducing a chimeric nucleic acid comprising  
CC a first sequence to regulate transcription, a second DNA sequence  
CC encoding a fusion polypeptide, comprising a sequence encoding an oleosin  
CC gene and sequence encoding (I) and a third sequence encoding a  
CC termination region functional in the host cell and growing the host cell  
CC to produce a fusion polypeptide. The method or (I) is useful for  
CC expression of a thioresdoxin or thioresdoxin reductase by a host cell. This  
CC sequence encodes a oleosin oil body protein fused to Arabidopsis thaliana  
CC thioresdoxin gene controlled by a phaseolin promoter and phaseolin  
CC terminator sequence.

SQ Sequence 3888 BP; 1254 A; 722 C; 613 G; 1299 T; 0 other;

Alignment Scores:

Pred. No.: 6.26e-45 Length: 3888  
Score: 411.00 Matches: 74  
Percent Similarity: 81.90% Conservative: 21  
Best Local Similarity: 63.79% Mismatches: 21  
Query Match: 66.50% Indels: 0  
DB: 24 Gaps: 0

US-09-786-715-8 (1-123) x ABSS3097 (1-3888)

QY 5 GluGluGluGluValAlIleGlyValAlHisThrValAspGluTrpLysLeuGlnLeuGlnAsn 24

DB 1564 GAAGAAGACAAAGATGCGCTGCGCACACCGTTGAGACATGGAAGACGAGCTTCAGAG 1623

QY 25 AlaLysAspSerLysLysLeuIleValAlAspPheThrAlaSerTrpCysGlyProCys 44

DB 1624 GCTAATGATCAAAAGCTGCTGCTGATTTCAAGCGCTCTGCTGCTGCTGCTGCTGCT 1683

QY 45 ArgPheMetAlaProValLeuAlaGluIleAlaLysLysThrProGluLeuIlePheLeu 64

DB 1684 CGTTTCATCGCTCAATCTTGTGATTTGGCTAAGAACTTCTTAACGTCCTTTCTCTC 1743

QY 65 LysValAspValAspGluValAlArgProValAlaGluGluTrpSerIleGluAlaMetPro 84

DB 1744 AAGCTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1803

QY 85 ThrPheLeuPheLeuLysAspGlyLysLeuIleValAlAspLysValAlGlyAlaSerLysAsp 104

Db 1804 ACCTTCATGTTTTGGAAGAGGAGACATTTTGGACAAAGTTGTTGGAGCCCAAGAACAT 18633  
105 ASpleuGlnAlaThrIleAlaIysHisAlaSerAlaValAlaAla 120  
:::111  
Db 1864 GAGCTTCAGCTCCTACCATTTGCCAACAACCTTGCTATGGCGATACAGCT 1911  
RESULT 9  
ID ABBN89581 standard; DNA; 3888 BP.  
XX ABBN89581;  
AC ABBN89581;  
XX ABBN89581;  
DT 06-SEP-2002 (first entry)  
XX  
DE Phaseolin promoter-Trxh oleosin-phaseolin terminator DNA SEQ:19.  
XX  
KW Multimeric protein; redox protein; thioredoxin; thioredoxin reductase;  
KW oil body; ophthalmological; antidiabetic; cytostatic; antiporiatic;  
KW vasotrophic; vulnerrary; antibacterial; immunosuppressive; antiulcer;  
KW food product; milk; wheat; oxidative stress; cataract; diabetes;  
KW chronic obstructive pulmonary disease; envenomation; psoriasis; sepsis;  
KW bronchiolupulmonary disease; malignancy; reperfusion injury; wound healing;  
KW gastro intestinal bleeding; intestinal bowel disease; ulcer;  
KW gastro oesophagael reflux disease; gene; ds.  
XX  
XX Arabidopsis sp.  
OS Phaseolus vulgaris.  
XX  
XX WO200250289-A1.  
PN 27-JUN-2002.  
XX  
XX 19-DEC-2001; 2001WO-US50240.  
PF 19-DEC-2000; 2000US-0742900.  
PR 05-JUL-2001; 2001US-302885P.  
ER 04-DEC-2001; 2001US-0006038.  
XX  
XX (SEMB-) SEMBIOSYS GENETICS INC.  
PA (SYGN) SYNGENTA PARTICIPATIONS AG.  
XX  
XX Van Rooijen G, Deckers H, Heifetz PB, Briggs SP, Dalmia BK;  
PI Del Val G, Zaplachinski S, Moloney M;  
XX  
XX WPI: 2002-508806/54.  
DR P-PSDB; ABB60683.  
XX  
XX  
XX The present invention describes a method (M1) for producing an oil body  
XX associated with a recombinant multimeric protein complex (MPC). M1  
XX comprises producing in a cell comprising oil bodies a first and second  
XX recombinant polypeptide (P1, P2), where P1 is capable of associating  
XX with P2 to form the MPC and associating the complex with an occlusion  
XX body (OB) through an OB-targeting-protein capable of associating with OB  
XX and P1. M1 is useful for producing an oil body associated with a  
XX recombinant MPC. The oil bodies are further formulated for use in the  
XX preparation of a food product such as milk or wheat based food product,  
XX personal care product which reduces the oxidative stress on the surface  
XX area of the human body or used to lighten the skin, or a pharmaceutical  
XX composition used to treat chronic obstructive pulmonary disease (COPD),  
XX cataracts, diabetes, envenomation, bronchiolupulmonary disease, psoriasis,  
XX malignancies, reperfusion injury, wound healing, sepsis, gastro  
XX intestinal (GI) bleeding, intestinal bowel disease (IBD), ulcers, GERD  
XX (gastro oesophagael reflux disease). ABBN89569 to ABBN89593 and ABB60677  
XX to ABB60964 represent sequence given in the exemplification of the  
XX present invention.

xx	Sequence	3888 BP; 1254 A; 722 C; 613 G; 1299 T; 0 other;
 Alignment Scores:		
Pred. No.:	6,266-45	Length: 3888
Score:	411.00	Matches: 74
Percent Similarity:	81.90%	Conservative: 21
'Best Local Similarity:	63.79%	Mismatches: 21
Query Match:	66.50%	Indels: 0
DB:	24	Gaps: 0
 US-09-786-715-8 (1-123) x ABN89581 (1-3888)		
OY	5	GlucgluglignValllleGlyValHlsThryValAspgIurTrplyLseuGlnLeuGlnasn 24      .....
Db	1564	GAAGAGACAAAGTATGCCTCGCCACACCCTGGAGACATGCAAGCAGACTTCACAG 16233      .....
OY	25	AlalysAsperLylysLeuIlleValValAspherHraIsertPrCygLyProCys 44     .....
Db	1624	GCTAATGATCCAAAACCTGTGGTGTTATTTACAGGGCTTCTGGTGAGCACATGT 16833     .....
OY	45	ArgPheMetAlAProValLeuAlleuAlagluIleAlaylsyThrProgluLeuIllePheLeu 64      .....
Db	1684	CgTTTCATCGCTCCATTCTTCTTGCGATTTGGCTTAAGAACACTTCTTAACGTCTTTCTC 17433      .....
OY	65	LysValAspValAspgLuValArgProValAlaglUutyrSerIlleguAlaMetPro 84      .....
Db	1744	AAGGTGATACTGATGAATTAAGTCGGTGCAGATGATGGCGATACGGCATGCCA 18033      .....
OY	85	ThrPheLeuPheLeuLysAspClgLuIlleValAsplysValAlglyAlaserLyasp 104      .....
Db	1804	ACCTTCAGTGTTTTGAGGAAGGAGAAATTTGGACAAAGTTGTGGAGCCAGAAAGAT 18633      .....
OY	105	AspleuGlnAlatrrlleAlalyshIsaseraIlaValAlaIala 120 :::     .....
Db	1864	GAGCTTAGCTACCATTTGCCAAACACTTGGCTATGGCGGATACAGCT 1911 :::     .....
 RESULT 10		
ID	ABN89586	standard; DNA; 4935 BP.
XX	ABN89586;	
AC		
XX		
DT	06-SEP-2002 (first entry)	
XX		
DE	Promoter-Oleosin thioredoxin reductase-linker-thioredoxin-terminator #33.	
XX		
KW	Multimeric protein; redox protein; thioredoxin; thioredoxin reductase;	
KW	oil body; ophthalmological; antidiabetic; cytosolic; antiportatic;	
KW	vasotrophic; vulnerable; antibacterial; immunosuppressive; anticancer;	
KW	food product; milk; wheat; oxidative stress; cataract; diabetes;	
KW	chronic obstructive pulmonary disease; envenomation; psoriasis; sepsis;	
KW	bronchiopulmonary disease; malignancy; reperfusion injury; wound healing;	
KW	gastro intestinal bleeding; intestinal bowel disease; ulcer;	
KW	gastro oesophageal reflux disease; gene; ds.	
XX		
OS	Arabidopsis sp.	
OS	Phaseolus vulgaris.	
XX		
PN	MO200250289-A1.	
XX		
PD	27-JUN-2002.	
XX		
PF	19-DEC-2001; 2001WO-US50240.	
XX		
PR	19-DEC-2000; 2000US-0742900.	
PR	05-JUL-2001; 2001US-302885P.	
PR	04-DEC-2001; 2001US-0006038.	
XX		
PA	(SEMB-) SEMBIOSYS GENETICS INC.	
PA	(SYGN) SYNGENTA PARTICIPATIONS AG.	
XX		
XX	Van Rooijen G., Deckers H., Helfetz PB., Briggs SP., Dalmia BK.	





PR 23-APR-1999; 99US-0130891.  
PR 28-APR-1999; 99US-0131449.  
PR 30-APR-1999; 99US-0132048.  
PR 30-APR-1999; 99US-0132407.  
PR 04-MAY-1999; 99US-0132484.  
PR 05-MAY-1999; 99US-0132485.  
PR 06-MAY-1999; 99US-0132486.  
PR 07-MAY-1999; 99US-0132487.  
PR 11-MAY-1999; 99US-0134256.  
PR 14-MAY-1999; 99US-0134218.  
PR 14-MAY-1999; 99US-0134219.  
PR 14-MAY-1999; 99US-0134221.  
PR 14-MAY-1999; 99US-0134370.  
PR 18-MAY-1999; 99US-01343768.  
PR 19-MAY-1999; 99US-0134941.  
PR 20-MAY-1999; 99US-0135124.  
PR 21-MAY-1999; 99US-0135353.  
PR 24-MAY-1999; 99US-0135629.  
PR 25-MAY-1999; 99US-0136021.  
PR 27-MAY-1999; 99US-0136392.  
PR 28-MAY-1999; 99US-0136782.  
PR 01-JUN-1999; 99US-0137222.  
PR 03-JUN-1999; 99US-0137528.  
PR 04-JUN-1999; 99US-0137502.  
PR 07-JUN-1999; 99US-0137724.  
PR 08-JUN-1999; 99US-0138094.  
PR 10-JUN-1999; 99US-0138540.  
PR 10-JUN-1999; 99US-0138847.  
PR 14-JUN-1999; 99US-0138119.  
PR 16-JUN-1999; 99US-0138452.  
PR 16-JUN-1999; 99US-0139453.  
PR 17-JUN-1999; 99US-0139492.  
PR 18-JUN-1999; 99US-0139454.  
PR 18-JUN-1999; 99US-0139455.  
PR 18-JUN-1999; 99US-0139456.  
PR 18-JUN-1999; 99US-0139457.  
PR 18-JUN-1999; 99US-0139458.  
PR 18-JUN-1999; 99US-0139459.  
PR 18-JUN-1999; 99US-0139460.  
PR 18-JUN-1999; 99US-0139461.  
PR 18-JUN-1999; 99US-0139462.  
PR 18-JUN-1999; 99US-0139463.  
PR 18-JUN-1999; 99US-0139750.  
PR 18-JUN-1999; 99US-0139763.  
PR 21-JUN-1999; 99US-0139817.  
PR 22-JUN-1999; 99US-0139899.  
PR 23-JUN-1999; 99US-0140353.  
PR 23-JUN-1999; 99US-0140354.  
PR 24-JUN-1999; 99US-0140695.  
PR 28-JUN-1999; 99US-0140823.  
PR 29-JUN-1999; 99US-0140991.  
PR 30-JUN-1999; 99US-0141287.  
PR 01-JUL-1999; 99US-0141842.  
PR 01-JUL-1999; 99US-0142154.  
PR 02-JUL-1999; 99US-0142055.  
PR 06-JUL-1999; 99US-0142390.  
PR 08-JUL-1999; 99US-0142803.  
PR 09-JUL-1999; 99US-0142920.  
PR 12-JUL-1999; 99US-0142977.  
PR 13-JUL-1999; 99US-0143542.  
PR 14-JUL-1999; 99US-0143624.  
PR 15-JUL-1999; 99US-0144005.  
PR 16-JUL-1999; 99US-0144085.  
PR 16-JUL-1999; 99US-0144086.  
PR 19-JUL-1999; 99US-0144325.  
PR 19-JUL-1999; 99US-0144331.  
PR 19-JUL-1999; 99US-0144332.  
PR 19-JUL-1999; 99US-0144333.  
PR 19-JUL-1999; 99US-0144334.  
PR 19-JUL-1999; 99US-0144335.  
PR 20-JUL-1999; 99US-0144352.  
PR 20-JUL-1999; 99US-0144632.  
  
PR 20-JUL-1999; 99US-0144884.  
PR 21-JUL-1999; 99US-0144814.  
PR 21-JUL-1999; 99US-0144866.  
PR 21-JUL-1999; 99US-0145088.  
PR 22-JUL-1999; 99US-0145085.  
PR 22-JUL-1999; 99US-0145087.  
PR 22-JUL-1999; 99US-0145089.  
PR 22-JUL-1999; 99US-0145192.  
PR 23-JUL-1999; 99US-0145145.  
PR 23-JUL-1999; 99US-0145218.  
PR 23-JUL-1999; 99US-0145224.  
PR 26-JUL-1999; 99US-0145276.  
PR 27-JUL-1999; 99US-0145913.  
PR 27-JUL-1999; 99US-0145918.  
PR 27-JUL-1999; 99US-0145919.  
PR 28-JUL-1999; 99US-0145951.  
PR 02-AUG-1999; 99US-0146386.  
PR 02-AUG-1999; 99US-0146388.  
PR 02-AUG-1999; 99US-0146389.  
PR 03-AUG-1999; 99US-0147038.  
PR 04-AUG-1999; 99US-0147204.  
PR 04-AUG-1999; 99US-0147302.  
PR 05-AUG-1999; 99US-0147192.  
PR 05-AUG-1999; 99US-0147260.  
PR 06-AUG-1999; 99US-0147303.  
PR 06-AUG-1999; 99US-0147416.  
PR 09-AUG-1999; 99US-0147493.  
PR 09-AUG-1999; 99US-0147935.  
PR 10-AUG-1999; 99US-0148171.  
PR 11-AUG-1999; 99US-0148319.  
PR 12-AUG-1999; 99US-0148341.  
PR 13-AUG-1999; 99US-0148565.  
PR 13-AUG-1999; 99US-0148684.  
PR 16-AUG-1999; 99US-0149368.  
PR 17-AUG-1999; 99US-0149175.  
PR 18-AUG-1999; 99US-0149426.  
PR 20-AUG-1999; 99US-0149722.  
PR 20-AUG-1999; 99US-0149723.  
PR 20-AUG-1999; 99US-0149929.  
PR 23-AUG-1999; 99US-0149902.  
PR 23-AUG-1999; 99US-0149930.  
PR 23-AUG-1999; 99US-0150566.  
PR 23-AUG-1999; 99US-0150567.  
PR 26-AUG-1999; 99US-0150884.  
PR 27-AUG-1999; 99US-0151065.  
PR 27-AUG-1999; 99US-0151066.  
PR 27-AUG-1999; 99US-0151080.  
PR 30-AUG-1999; 99US-0151303.  
PR 31-AUG-1999; 99US-0151348.  
PR 01-SEP-1999; 99US-0151930.  
PR 07-SEP-1999; 99US-0152363.  
PR 10-SEP-1999; 99US-0153070.  
PR 13-SEP-1999; 99US-0153758.  
PR 15-SEP-1999; 99US-0154018.  
PR 16-SEP-1999; 99US-0154039.  
PR 20-SEP-1999; 99US-0154779.  
PR 22-SEP-1999; 99US-0155139.  
PR 23-SEP-1999; 99US-0155486.  
PR 24-SEP-1999; 99US-0155659.  
PR 28-SEP-1999; 99US-0156456.  
PR 29-SEP-1999; 99US-0156596.  
PR 04-OCT-1999; 99US-0157117.  
PR 05-OCT-1999; 99US-0157753.  
PR 06-OCT-1999; 99US-0157865.  
PR 07-OCT-1999; 99US-0158029.  
PR 08-OCT-1999; 99US-0158232.  
PR 12-OCT-1999; 99US-0158369.  
PR 13-OCT-1999; 99US-0159293.  
PR 13-OCT-1999; 99US-0159294.  
PR 13-OCT-1999; 99US-0159329.  
PR 14-OCT-1999; 99US-0159329.  
PR 14-OCT-1999; 99US-0159330.  
PR 14-OCT-1999; 99US-0159331.  
PR 14-OCT-1999; 99US-0159637.



PR 14-OCT-1999; 99US-0159638.  
PR 18-OCT-1999; 99US-0159584.  
PR 21-OCT-1999; 99US-0160741.  
PR 21-OCT-1999; 99US-0160767.  
PR 21-OCT-1999; 99US-0160768.  
PR 21-OCT-1999; 99US-0160770.  
PR 21-OCT-1999; 99US-0160814.  
PR 21-OCT-1999; 99US-0160815.  
PR 22-OCT-1999; 99US-0160880.  
PR 22-OCT-1999; 99US-0160981.  
PR 22-OCT-1999; 99US-0160989.  
PR 25-OCT-1999; 99US-0161404.  
PR 25-OCT-1999; 99US-0161405.  
PR 25-OCT-1999; 99US-0161406.  
PR 26-OCT-1999; 99US-0161359.  
PR 26-OCT-1999; 99US-0161360.  
PR 26-OCT-1999; 99US-0161361.  
PR 28-OCT-1999; 99US-0161920.  
PR 28-OCT-1999; 99US-0161992.  
PR 28-OCT-1999; 99US-0161993.  
PR 29-OCT-1999; 99US-0162142.

## Alignment Scores:

Pred. No.: 4.45e-46  
Score: 410.00  
Percent Similarity: 85.328  
Best Local Similarity: 66.978  
Query Match: 66.348  
DB: 21

Length: 470  
Matches: 73  
Conservative: 20  
Mismatches: 16  
Indels: 0  
Gaps: 0

US-09-786-715-8 (1-123) x AAC37781 (1-470)

QY 5 Glllgluglyglvalilleglyvalihstrvalaspglutrlpysleuglnleuglnasn 24  
DB 104 GAGAGAGGACAGTAGTGCCTGCCACACGTTGAGCATGAGCAGAGCAGCTTCAGAG 163  
QY 25 AlalysAspSerLysLysLeuileValaspPhetThrAlaSerTrpCysglProCys 44  
DB 164 GCTTATGATCAACAACTTGTGGGTGATTCACGGCTTCCTGGTGGACCATAT 223  
QY 45 ArgpHeMetaIProValleuIaIleuIaIalysLysThProgluLeuilepHeleu 64  
DB 224 CTTTCATCGCTCCATCTTCTGCTGATTTGGCTAGAAACTTCCTAACGCTTTTCC 283  
QY 65 LysValaspValaspGlValarProValaIaIagluIuTySerIleGluaIaMetPro 84  
DB 284 AAGGTGATGATGATGAATTAAGTCGTCGCAAGTGGCGCATGACGCGATGCCA 343  
QY 85 ThrpHeuPheLeuLysaspLysaspLysLysValaspLysValaIaIaSerLysasp 104  
DB 344 ACCCTTCATGTTTTCAGAGAGGAGGAGATTTTGGACAAAGTTGTTGGAGCCAGAAAGAT 403  
QY 105 AspleuglAlaIThrIleAlaLysHls 113  
DB 404 GAGCTTCAGTACCATTCACCAACAC 430

## RESULT 14

AAC51522 standard; DNA; 561 BP.

AAC51522;

18-OCT-2000 (first entry)

DE Arabidopsis thaliana DNA fragment SEQ ID NO: 66832.

XX Hybridisation assay; genetic mapping; gene expression control;

KW protein identification; signal transduction pathway;

XX metabolic pathway; promoter; termination sequence; ss.

OS Arabidopsis thaliana.

PN EPI033405-A2.

XX 06-SEP-2000.  
PD 25-FEB-2000; 2000EP-0301439.  
XX 25-FEB-1999; 99US-0121825.  
PR 05-MAR-1999; 99US-0123180.  
PR 09-MAR-1999; 99US-0123548.  
PR 23-MAR-1999; 99US-0125788.  
PR 25-MAR-1999; 99US-0126264.  
PR 29-MAR-1999; 99US-0126785.  
PR 01-APR-1999; 99US-0127462.  
PR 06-APR-1999; 99US-0128234.  
PR 08-APR-1999; 99US-0128714.  
PR 16-APR-1999; 99US-0129645.  
PR 19-APR-1999; 99US-0130077.  
PR 21-APR-1999; 99US-0130449.  
PR 23-APR-1999; 99US-0130510.  
PR 23-APR-1999; 99US-0130891.  
PR 28-APR-1999; 99US-0131449.  
PR 30-APR-1999; 99US-0132048.  
PR 30-APR-1999; 99US-0132407.  
PR 04-MAY-1999; 99US-0132484.  
PR 05-MAY-1999; 99US-0132485.  
PR 06-MAY-1999; 99US-0132486.  
PR 07-MAY-1999; 99US-0132487.  
PR 07-MAY-1999; 99US-0132863.  
PR 11-MAY-1999; 99US-0134256.  
PR 14-MAY-1999; 99US-0134218.  
PR 14-MAY-1999; 99US-0134219.  
PR 14-MAY-1999; 99US-0134421.  
PR 14-MAY-1999; 99US-0134422.  
PR 14-MAY-1999; 99US-0134423.  
PR 18-MAY-1999; 99US-0134768.  
PR 19-MAY-1999; 99US-0134941.  
PR 20-MAY-1999; 99US-0135124.  
PR 21-MAY-1999; 99US-0135353.  
PR 24-MAY-1999; 99US-0135629.  
PR 25-MAY-1999; 99US-0136021.  
PR 27-MAY-1999; 99US-0136392.  
PR 28-MAY-1999; 99US-0136782.  
PR 01-JUN-1999; 99US-0137222.  
PR 03-JUN-1999; 99US-0137528.  
PR 04-JUN-1999; 99US-0137502.  
PR 07-JUN-1999; 99US-0137724.  
PR 08-JUN-1999; 99US-0138094.  
PR 10-JUN-1999; 99US-0138540.  
PR 10-JUN-1999; 99US-0138847.  
PR 14-JUN-1999; 99US-0139119.  
PR 16-JUN-1999; 99US-0139452.  
PR 16-JUN-1999; 99US-0139453.  
PR 17-JUN-1999; 99US-0139492.  
PR 18-JUN-1999; 99US-0139454.  
PR 18-JUN-1999; 99US-0139455.  
PR 18-JUN-1999; 99US-0139456.  
PR 18-JUN-1999; 99US-0139457.  
PR 18-JUN-1999; 99US-0139458.  
PR 18-JUN-1999; 99US-0139459.  
PR 21-JUN-1999; 99US-0139817.  
PR 22-JUN-1999; 99US-0139899.  
PR 23-JUN-1999; 99US-0140353.  
PR 23-JUN-1999; 99US-0140354.  
PR 24-JUN-1999; 99US-0140695.  
PR 28-JUN-1999; 99US-0140823.  
PR 29-JUN-1999; 99US-0140991.  
PR 30-JUN-1999; 99US-0141287.  
PR 01-JUL-1999; 99US-0141842.  
PR 01-JUL-1999; 99US-0142154.

PR 02-JUL-1999; 99US-0142055.  
PR 06-JUL-1999; 99US-0142390.  
PR 08-JUL-1999; 99US-0142803.  
PR 09-JUL-1999; 99US-0142920.  
PR 12-JUL-1999; 99US-0142977.  
PR 13-JUL-1999; 99US-0143542.  
PR 14-JUL-1999; 99US-0143624.  
PR 15-JUL-1999; 99US-0144005.  
PR 16-JUL-1999; 99US-0144085.  
PR 16-JUL-1999; 99US-0144086.  
PR 19-JUL-1999; 99US-0144325.  
PR 19-JUL-1999; 99US-0144331.  
PR 19-JUL-1999; 99US-0144332.  
PR 19-JUL-1999; 99US-0144333.  
PR 19-JUL-1999; 99US-0144334.  
PR 19-JUL-1999; 99US-0144335.  
PR 20-JUL-1999; 99US-0144352.  
PR 20-JUL-1999; 99US-0144632.  
PR 20-JUL-1999; 99US-0144884.  
PR 21-JUL-1999; 99US-0144814.  
PR 21-JUL-1999; 99US-0145086.  
PR 21-JUL-1999; 99US-0145088.  
PR 22-JUL-1999; 99US-0145085.  
PR 22-JUL-1999; 99US-0145087.  
PR 22-JUL-1999; 99US-0145089.  
PR 22-JUL-1999; 99US-0145192.  
PR 23-JUL-1999; 99US-0145145.  
PR 23-JUL-1999; 99US-0145218.  
PR 23-JUL-1999; 99US-0145224.  
PR 26-JUL-1999; 99US-0145276.  
PR 27-JUL-1999; 99US-0145913.  
PR 27-JUL-1999; 99US-0145918.  
PR 27-JUL-1999; 99US-0145919.  
PR 28-JUL-1999; 99US-0145951.  
PR 02-AUG-1999; 99US-0146386.  
PR 02-AUG-1999; 99US-0146388.  
PR 02-AUG-1999; 99US-0146389.  
PR 03-AUG-1999; 99US-0147038.  
PR 04-AUG-1999; 99US-0147204.  
PR 04-AUG-1999; 99US-0147302.  
PR 05-AUG-1999; 99US-0147192.  
PR 05-AUG-1999; 99US-0147260.  
PR 06-AUG-1999; 99US-0147303.  
PR 06-AUG-1999; 99US-0147416.  
PR 09-AUG-1999; 99US-0147493.  
PR 09-AUG-1999; 99US-0147935.  
PR 10-AUG-1999; 99US-0148171.  
PR 11-AUG-1999; 99US-0148319.  
PR 12-AUG-1999; 99US-0148341.  
PR 13-AUG-1999; 99US-0148565.  
PR 13-AUG-1999; 99US-0148684.  
PR 16-AUG-1999; 99US-0149368.  
PR 17-AUG-1999; 99US-0149175.  
PR 18-AUG-1999; 99US-0149426.  
PR 20-AUG-1999; 99US-0149722.  
PR 20-AUG-1999; 99US-0149723.  
PR 20-AUG-1999; 99US-0149929.  
PR 23-AUG-1999; 99US-0149902.  
PR 23-AUG-1999; 99US-0149930.  
PR 25-AUG-1999; 99US-0150566.  
PR 26-AUG-1999; 99US-0150884.  
PR 26-AUG-1999; 99US-0151065.  
PR 27-AUG-1999; 99US-0151066.  
PR 27-AUG-1999; 99US-0151080.  
PR 30-AUG-1999; 99US-0151303.  
PR 31-AUG-1999; 99US-0151438.  
PR 01-SEP-1999; 99US-0151930.  
PR 07-SEP-1999; 99US-0152363.  
PR 10-SEP-1999; 99US-0153070.  
PR 13-SEP-1999; 99US-0153758.  
PR 15-SEP-1999; 99US-0154018.  
PR 16-SEP-1999; 99US-0154039.  
PR 20-SEP-1999; 99US-0154779.

PR 22-SEP-1999; 99US-0155139.  
PR 23-SEP-1999; 99US-0155486.  
PR 24-SEP-1999; 99US-0155659.  
PR 28-SEP-1999; 99US-0156458.  
PR 29-SEP-1999; 99US-0156596.  
PR 04-OCT-1999; 99US-0157117.  
PR 05-OCT-1999; 99US-0157173.  
PR 06-OCT-1999; 99US-0157865.  
PR 07-OCT-1999; 99US-0158029.  
PR 08-OCT-1999; 99US-0158232.  
PR 12-OCT-1999; 99US-0158369.  
PR 13-OCT-1999; 99US-0158293.  
PR 13-OCT-1999; 99US-0159293.  
PR 13-OCT-1999; 99US-0159295.  
PR 14-OCT-1999; 99US-0159329.  
PR 14-OCT-1999; 99US-0159330.  
PR 14-OCT-1999; 99US-0159331.  
PR 14-OCT-1999; 99US-0159637.  
PR 14-OCT-1999; 99US-0159638.  
PR 18-OCT-1999; 99US-0159584.  
PR 21-OCT-1999; 99US-0160741.  
PR 21-OCT-1999; 99US-0160767.  
PR 21-OCT-1999; 99US-0160768.  
PR 21-OCT-1999; 99US-0160770.  
PR 21-OCT-1999; 99US-0160815.  
PR 21-OCT-1999; 99US-0160815.  
PR 22-OCT-1999; 99US-0160980.  
PR 22-OCT-1999; 99US-0160981.  
PR 22-OCT-1999; 99US-0160989.  
PR 25-OCT-1999; 99US-0161404.  
PR 25-OCT-1999; 99US-0161405.  
PR 25-OCT-1999; 99US-0161406.  
PR 26-OCT-1999; 99US-0161359.  
PR 26-OCT-1999; 99US-0161360.  
PR 26-OCT-1999; 99US-0161361.  
PR 28-OCT-1999; 99US-0161920.  
PR 28-OCT-1999; 99US-0161992.  
PR 28-OCT-1999; 99US-0161993.  
PR 29-OCT-1999; 99US-0162142.

## Alignment Scores:

Pred. No.: 5,7e-46 Length: 561  
Score: 410.00 Matches: 73  
Percent Similarity: 85.32% Conservative: 20  
Best Local Similarity: 66.97% Mismatches: 16  
Query Match: 66.34% Indels: 0  
DB: 21 Gaps: 0

US-09-786-715-8 (1-123) x AAC51522 (1-561)

QY 5 GIUGLUGLGNValIIeGLyValIIeIstHrValAspGLuTrpLysLeuGlnAsn 24  
|||||  
DB 103 GANAGAGACAGATGATCGCTGCCACACCGTTGACATGGAAGCAAGCTTCAGAG 162  
|||  
QY 25 AAlAspSerLysLysLeuIIeValValAspHrThraLAserrPcysGlyProCys 44  
|||  
DB 163 GCTATGATCAACAAACTCTGTGGTGGTGGATTTTCACGGCTCTTGGTGGACATGT 222  
|||  
QY 45 ArgPheMetLarProValIIeAlaGluIIeAlaLysLysHrProGluLeuIIePheLeu 64  
|||||  
DB 223 CGTTTCATCGCTCCATCTTCTGCTGATTTGGCTAAGAACTCTCAAGTCTTTCCTC 282  
|||||  
QY 65 LysValAspValAspGluValArgProValAlaGluLutHrSerIIeGluAlaMetPro 84  
|||||  
DB 283 AAGTTTATCTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 342  
|||||  
QY 85 ThrPheLeuPheLeuLysAspGlyGluIIeValAspLysValAlaGlyAlaSerLysAsp 104  
|||||  
DB 343 ACCTTCATGTTTGTGAAGGAAGGAAGATTTTGACAAAGTGTGTGAGCAAGAAAGAT 402  
|||||  
QY 105 AspleuGlnAlaHrIIeAlaLysHis 113  
|||||  
DB 403 GAGCTTCAGTCTACCATTTGCCAAACAC 429

RESULT 15  
AAC34121  
ID AAC34121 standard; DNA: 563 BP.  
XX  
AC AAC34121;  
XX  
DT 17-OCT-2000 (first entry)  
XX  
DE Arabidopsis thaliana DNA fragment SEQ ID NO: 5524.  
XX  
KM Hybridisation assay; genetic mapping; gene expression control;  
KM protein identification; signal transduction pathway;  
KM metabolic pathway; promoter; termination sequence; ss.  
XX  
OS Arabidopsis thaliana.  
XX  
PN EP1033405-A2.  
XX  
PD 06-SEP-2000.  
XX  
PF 25-FEB-2000; 2000EP-0301439.  
XX  
PR 25-FEB-1999; 99US-0121825.  
PR 05-MAR-1999; 99US-0123180.  
PR 09-MAR-1999; 99US-0123548.  
PR 23-MAR-1999; 99US-0125788.  
PR 25-MAR-1999; 99US-0126264.  
PR 29-MAR-1999; 99US-0126785.  
PR 01-APR-1999; 99US-0127462.  
PR 06-APR-1999; 99US-0128234.  
PR 08-APR-1999; 99US-0128714.  
PR 16-APR-1999; 99US-0129645.  
PR 19-APR-1999; 99US-0130077.  
PR 21-APR-1999; 99US-0130449.  
PR 23-APR-1999; 99US-0130510.  
PR 23-APR-1999; 99US-0130891.  
PR 28-APR-1999; 99US-0131449.  
PR 30-APR-1999; 99US-0132048.  
PR 30-APR-1999; 99US-0132407.  
PR 04-MAY-1999; 99US-0132484.  
PR 05-MAY-1999; 99US-0132485.  
PR 06-MAY-1999; 99US-0132486.  
PR 07-MAY-1999; 99US-0132487.  
PR 11-MAY-1999; 99US-0132863.  
PR 11-MAY-1999; 99US-0134256.  
PR 14-MAY-1999; 99US-0134218.  
PR 14-MAY-1999; 99US-0134219.  
PR 14-MAY-1999; 99US-0134221.  
PR 14-MAY-1999; 99US-0134370.  
PR 18-MAY-1999; 99US-0134768.  
PR 19-MAY-1999; 99US-0134941.  
PR 20-MAY-1999; 99US-0135124.  
PR 21-MAY-1999; 99US-0135353.  
PR 24-MAY-1999; 99US-0135629.  
PR 25-MAY-1999; 99US-0136021.  
PR 27-MAY-1999; 99US-0136392.  
PR 28-MAY-1999; 99US-0136782.  
PR 01-JUN-1999; 99US-0137222.  
PR 03-JUN-1999; 99US-0137528.  
PR 04-JUN-1999; 99US-0137502.  
PR 07-JUN-1999; 99US-0137724.  
PR 08-JUN-1999; 99US-0138094.  
PR 10-JUN-1999; 99US-0138540.  
PR 10-JUN-1999; 99US-0138847.  
PR 14-JUN-1999; 99US-0139119.  
PR 16-JUN-1999; 99US-0139452.  
PR 16-JUN-1999; 99US-0139453.  
PR 17-JUN-1999; 99US-0139492.  
PR 18-JUN-1999; 99US-0139454.  
PR 18-JUN-1999; 99US-0139455.  
PR 18-JUN-1999; 99US-0139456.  
PR 18-JUN-1999; 99US-0139457.

PR 18-JUN-1999; 99US-0139458.  
PR 18-JUN-1999; 99US-0139459.  
PR 18-JUN-1999; 99US-0139460.  
PR 18-JUN-1999; 99US-0139461.  
PR 18-JUN-1999; 99US-0139462.  
PR 18-JUN-1999; 99US-0139463.  
PR 18-JUN-1999; 99US-0139750.  
PR 18-JUN-1999; 99US-0139753.  
PR 21-JUN-1999; 99US-0139817.  
PR 22-JUN-1999; 99US-0139899.  
PR 23-JUN-1999; 99US-0140353.  
PR 23-JUN-1999; 99US-0140354.  
PR 24-JUN-1999; 99US-0140695.  
PR 28-JUN-1999; 99US-0140823.  
PR 29-JUN-1999; 99US-0140991.  
PR 30-JUN-1999; 99US-0141287.  
PR 01-JUL-1999; 99US-0141842.  
PR 01-JUL-1999; 99US-0142154.  
PR 02-JUL-1999; 99US-0142055.  
PR 06-JUL-1999; 99US-0142390.  
PR 08-JUL-1999; 99US-0142803.  
PR 09-JUL-1999; 99US-0142920.  
PR 12-JUL-1999; 99US-0142977.  
PR 13-JUL-1999; 99US-0143542.  
PR 14-JUL-1999; 99US-0143624.  
PR 15-JUL-1999; 99US-0144005.  
PR 16-JUL-1999; 99US-0144085.  
PR 16-JUL-1999; 99US-0144086.  
PR 19-JUL-1999; 99US-0144325.  
PR 19-JUL-1999; 99US-0144331.  
PR 19-JUL-1999; 99US-0144332.  
PR 19-JUL-1999; 99US-0144333.  
PR 19-JUL-1999; 99US-0144334.  
PR 19-JUL-1999; 99US-0144335.  
PR 20-JUL-1999; 99US-0144352.  
PR 20-JUL-1999; 99US-0144632.  
PR 21-JUL-1999; 99US-0144884.  
PR 21-JUL-1999; 99US-0144814.  
PR 21-JUL-1999; 99US-0145086.  
PR 21-JUL-1999; 99US-0145088.  
PR 22-JUL-1999; 99US-0145085.  
PR 22-JUL-1999; 99US-0145087.  
PR 22-JUL-1999; 99US-0145089.  
PR 22-JUL-1999; 99US-0145192.  
PR 23-JUL-1999; 99US-0145145.  
PR 23-JUL-1999; 99US-0145218.  
PR 23-JUL-1999; 99US-0145224.  
PR 26-JUL-1999; 99US-0145276.  
PR 27-JUL-1999; 99US-0145293.  
PR 27-JUL-1999; 99US-0145918.  
PR 27-JUL-1999; 99US-0145919.  
PR 28-JUL-1999; 99US-0145951.  
PR 02-AUG-1999; 99US-0146386.  
PR 02-AUG-1999; 99US-0146388.  
PR 02-AUG-1999; 99US-0146389.  
PR 03-AUG-1999; 99US-0147308.  
PR 04-AUG-1999; 99US-0147204.  
PR 04-AUG-1999; 99US-0147302.  
PR 05-AUG-1999; 99US-0147192.  
PR 05-AUG-1999; 99US-0147260.  
PR 06-AUG-1999; 99US-0147303.  
PR 06-AUG-1999; 99US-0147416.  
PR 09-AUG-1999; 99US-0147493.  
PR 09-AUG-1999; 99US-0147935.  
PR 10-AUG-1999; 99US-0148171.  
PR 11-AUG-1999; 99US-0148319.  
PR 12-AUG-1999; 99US-0148341.  
PR 13-AUG-1999; 99US-0148565.  
PR 13-AUG-1999; 99US-0148684.  
PR 16-AUG-1999; 99US-0149368.  
PR 17-AUG-1999; 99US-0149175.  
PR 18-AUG-1999; 99US-0149426.  
PR 20-AUG-1999; 99US-0149722.

PR 20-AUG-1999; 99US-0149723.  
PR 20-AUG-1999; 99US-0149929.  
PR 23-AUG-1999; 99US-0149902.  
PR 23-AUG-1999; 99US-0149930.  
PR 25-AUG-1999; 99US-0150566.  
PR 26-AUG-1999; 99US-0150884.  
PR 27-AUG-1999; 99US-0151065.  
PR 27-AUG-1999; 99US-0151066.  
PR 27-AUG-1999; 99US-0151080.  
PR 30-AUG-1999; 99US-0151303.  
PR 31-AUG-1999; 99US-0151438.  
PR 01-SEP-1999; 99US-0151930.  
PR 07-SEP-1999; 99US-0152363.  
PR 10-SEP-1999; 99US-0153070.  
PR 13-SEP-1999; 99US-0153758.  
PR 15-SEP-1999; 99US-0154018.  
PR 16-SEP-1999; 99US-0154039.  
PR 20-SEP-1999; 99US-0154779.  
PR 22-SEP-1999; 99US-0155139.  
PR 23-SEP-1999; 99US-0155486.  
PR 24-SEP-1999; 99US-0155659.  
PR 28-SEP-1999; 99US-0156458.  
PR 29-SEP-1999; 99US-0156596.  
PR 04-OCT-1999; 99US-0157117.  
PR 05-OCT-1999; 99US-0157753.  
PR 06-OCT-1999; 99US-0157865.  
PR 07-OCT-1999; 99US-0158029.  
PR 08-OCT-1999; 99US-0158232.  
PR 12-OCT-1999; 99US-0158368.  
PR 13-OCT-1999; 99US-0159293.  
PR 13-OCT-1999; 99US-0159294.  
PR 13-OCT-1999; 99US-0159295.  
PR 14-OCT-1999; 99US-0159329.  
PR 14-OCT-1999; 99US-0159330.  
PR 14-OCT-1999; 99US-0159331.  
PR 14-OCT-1999; 99US-0159337.  
PR 14-OCT-1999; 99US-0159638.  
PR 18-OCT-1999; 99US-0159584.  
PR 21-OCT-1999; 99US-0160741.  
PR 21-OCT-1999; 99US-0160767.  
PR 21-OCT-1999; 99US-0160768.  
PR 21-OCT-1999; 99US-0160770.  
PR 21-OCT-1999; 99US-0160814.  
PR 21-OCT-1999; 99US-0160815.  
PR 22-OCT-1999; 99US-0160980.  
PR 22-OCT-1999; 99US-0160981.  
PR 22-OCT-1999; 99US-0160989.  
PR 25-OCT-1999; 99US-0161404.  
PR 25-OCT-1999; 99US-0161405.  
PR 25-OCT-1999; 99US-0161406.  
PR 26-OCT-1999; 99US-0161359.  
PR 26-OCT-1999; 99US-0161360.  
PR 26-OCT-1999; 99US-0161361.  
PR 28-OCT-1999; 99US-0161920.  
PR 28-OCT-1999; 99US-0161922.  
PR 28-OCT-1999; 99US-0161993.  
PR 29-OCT-1999; 99US-0162142.

## Alignment Scores:

Pred. No.: 5.73e-46  
Score: 410.00  
Percent Similarity: 85.32%  
Best Local Similarity: 66.97%  
Query Match: 66.34%  
DB: 21  
Length: 563  
Matches: 73  
Conservative: 20  
Mismatch: 16  
Indels: 0  
Gaps: 0

US-09-786-715-8 (1-123) x AAC34121 (1-563)

OY 5 GluGluGlyGlnValIleGlyValHisThrValAspGluTrpLysLeuGlnLeuGlnAsn 24  
Db 105 GAAAGAGCAAGCAAGTGCCTGCCACACCGTTGAGACATGGAAGACAGACAGCTTCAAGAG 164  
OY 25 AlAlAspSerLysLysLeuIleValValAspPheThrAlaSerTrpCysGlyProCys 44

DB 165 GCTATATGATCCAAACACTCTTGTGTGATTCACGGCTTGTGTGACCATGT 224  
OY 45 ArgPheMetAlaProValLeuAlaIleAlaLysLysThrProGluLeuIlePheLeu 64  
DB 225 CGTTTCATCGCCGCACTTCTTTCATTTGGCTAAGAAACCTTCACGCTTTCCTC 284  
OY 65 LysValAspValAspGluValArgProValAlaGluGluTrpSerIleGluAlaMetPro 84  
DB 285 AAGTTGATACGATGATGATGAGTCGGGCAAGTGTGGCCATACAGCCATGCCA 344  
OY 85 ThrPheLeuPheLeuLysAspGlyIleIleValAspLysValAlaGlyAlaSerLysAsp 104  
DB 345 ACCTTCATGTTTTCAGAGCAAGCAATTTGCAAAAGTTGTGAGCCAAAGAGAT 404  
OY 105 AspLeuGlnAlaThrIleAlaLysHis 113  
DB 405 GAGCTTCAGTCATCACCATTCGCCAAACAC 431

Search completed: August 17, 2003, 19:37:18  
Job time : 176.913 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - nucleic search, using frame\_plus.p2n model

Run on: August 17, 2003, 19:18:57 ; Search time 45.5839 Seconds  
(without alignments)  
1190.995 Million cell updates/sec

Title: US-09-786-715-8

Perfect score: 618

Sequence: 1 MAEEVGQYIGVHTDEMKL.....DDLQATIAKASAVNAASS 123

Scoring table: BLOSUM62  
Xgapop 10.0, Xgapext 0.5  
Ygapop 10.0, Ygapext 0.5  
Fgapop 6.0, Fgapext 7.0  
Delop 6.0, Delext 7.0

Searched: 569978 seqs, 220691566 residues

Total number of hits satisfying chosen parameters: 1139956

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Command line parameters:

-MODEL=frame\_plus.p2n.model -DEV=xlp  
-O=/cgn2\_1/uspt0.spool/US09786715/rnat\_11082003\_150515\_6083/app\_query.fasta\_1.1052  
-DB=Issued\_Patents\_NA -QFMT=fastap -SUFFIX=tni -MINMATCH=0.1 -LOOPEXT=0  
-LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blomsu62 -TRANS=humana0.cdi  
-LIST=45 -DOCALLIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15  
-MODE=LOCAL -OUTFMT=ptc -NORM=ext -HAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000  
-USER=US09786715.eccn\_1.175.etrnat\_11082003\_150515\_6083 -MCPu=6 -ICPU=3  
-NO\_MAMP -LARGEQUERY -NEG\_SCORES=0 -WAIT -DSPBLCK=100 -LONGLOG  
-DEV\_TIMEOUT=120 -WARN\_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6  
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

Issued\_Patents\_NA:\*  
1: /cgn2\_6/ptodata/1/ina/5A.COMB.seq:\*  
2: /cgn2\_6/ptodata/1/ina/5B.COMB.seq:\*  
3: /cgn2\_6/ptodata/1/ina/6A.COMB.seq:\*  
4: /cgn2\_6/ptodata/1/ina/6B.COMB.seq:\*  
5: /cgn2\_6/ptodata/1/ina/PCUTUS.COMB.seq:\*  
6: /cgn2\_6/ptodata/1/ina/Backfillseq1.seq:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	392	63.4	393	4	US-09-540-014-5
2	392	63.4	653	1	US-08-181-2715-103
3	392	63.4	653	1	US-08-449-3175-103
4	392	63.4	653	1	US-08-444-803-103
5	392	63.4	653	1	US-08-449-043-103
6	392	63.4	653	1	US-08-456-265A-103
7	392	63.4	653	1	US-08-455-416-103
8	392	63.4	653	1	US-08-455-244-103
9	392	63.4	653	1	US-08-454-876-103
10	392	63.4	653	2	US-08-457-364-103
11	392	63.4	653	2	US-08-456-262-103
12	392	63.4	653	2	US-08-456-240-103

13	392	63.4	653	2	US-08-455-736-103	Sequence 103, App
14	392	63.4	653	2	US-08-971-217-103	Sequence 103, App
15	392	63.4	653	3	US-09-350-600-103	Sequence 103, App
16	389	62.9	359	4	US-09-540-014-1	Sequence 1, Appl1
17	388	62.8	382	4	US-09-540-014-3	Sequence 3, Appl1
18	388	62.8	557	4	US-09-404-879A-88	Sequence 88, Appl1
19	388	62.8	557	4	US-09-338-933-88	Sequence 88, Appl1
20	388	62.8	557	4	US-09-215-681-88	Sequence 88, Appl1
21	388	62.8	557	4	US-09-215-681-88	Sequence 88, Appl1
22	388	62.8	557	4	US-09-215-681-88	Sequence 88, Appl1
23	388	62.8	557	4	US-09-215-681-88	Sequence 88, Appl1
24	388	62.8	557	4	US-09-215-681-88	Sequence 88, Appl1
25	388	62.8	557	4	US-09-215-681-88	Sequence 88, Appl1
26	388	62.8	557	4	US-09-215-681-88	Sequence 88, Appl1
27	388	62.8	557	4	US-09-215-681-88	Sequence 88, Appl1
28	388	62.8	557	4	US-09-215-681-88	Sequence 88, Appl1
29	388	62.8	557	4	US-09-215-681-88	Sequence 88, Appl1
30	388	62.8	557	4	US-09-215-681-88	Sequence 88, Appl1
31	388	62.8	557	4	US-09-215-681-88	Sequence 88, Appl1
32	388	62.8	557	4	US-09-215-681-88	Sequence 88, Appl1
33	388	62.8	557	4	US-09-215-681-88	Sequence 88, Appl1
34	388	62.8	557	4	US-09-215-681-88	Sequence 88, Appl1
35	388	62.8	557	4	US-09-215-681-88	Sequence 88, Appl1
36	388	62.8	557	4	US-09-215-681-88	Sequence 88, Appl1
37	388	62.8	557	4	US-09-215-681-88	Sequence 88, Appl1
38	388	62.8	557	4	US-09-215-681-88	Sequence 88, Appl1
39	388	62.8	557	4	US-09-215-681-88	Sequence 88, Appl1
40	388	62.8	557	4	US-09-215-681-88	Sequence 88, Appl1
41	388	62.8	557	4	US-09-215-681-88	Sequence 88, Appl1
42	388	62.8	557	4	US-09-215-681-88	Sequence 88, Appl1
43	388	62.8	557	4	US-09-215-681-88	Sequence 88, Appl1
44	388	62.8	557	4	US-09-215-681-88	Sequence 88, Appl1
45	388	62.8	557	4	US-09-215-681-88	Sequence 88, Appl1

#### ALIGNMENTS

RESULT 1  
US-09-540-014-5  
; Sequence 5, Application US/09540014  
; Patent No. 6380372  
; GENERAL INFORMATION:  
; APPLICANT: Cho, Myeong-Je  
; APPLICANT: Del Val, Greg  
; APPLICANT: Calliau, Maxime  
; APPLICANT: Lemauz, Peggy G.  
; APPLICANT: Buchanan, Bob B.  
; TITLE OF INVENTION: Barley Gene for Thioredoxin and  
; FILE REFERENCE: 2001-0701.30  
; CURRENT APPLICATION NUMBER: US/09/540,014  
; PRIOR FILING DATE: 2000-03-31  
; PRIOR APPLICATION NUMBER: US 60/127,198  
; PRIOR FILING DATE: 1999-03-31  
; PRIOR APPLICATION NUMBER: US 60/169,162  
; PRIOR FILING DATE: 1999-12-06  
; PRIOR APPLICATION NUMBER: US 60/177,740  
; PRIOR FILING DATE: 2000-01-21  
; PRIOR APPLICATION NUMBER: US 60/177,739  
; PRIOR FILING DATE: 2000-01-21  
; NUMBER OF SEQ ID NOS: 51  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 5  
; LENGTH: 393  
; TYPE: DNA  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: wheat thioredoxin h cDNA  
US-09-540-014-5

Alignment Scores: 6.18e-46 Length: 393  
Pred. No.: 392.00 Matches: 65  
Score:

Percent Similarity: 82.61% Conservative: 30  
Best Local Similarity: 56.52% Mismatches: 20  
Query Match: 63.43% Indels: 0  
DB: 4 Gaps: 0

US-09-786-715-8 (1-123) x US-09-540-014-5 (1-393)

Oy 2 AlaGluValGluGluGluValIleGlyValHisThrValAspGluTrpLeuGln 21  
Db 46 GCGCGCGTGGGGGGGGAGGTGATCTCCGTCACACCGCTGGAGCATGACATGCGAG 105  
Oy 22 LeuGlnAsnAlaLysAspSerLysLeuIleValValAspPheThrAlaSerTrpCys 41  
Db 106 ATCAGAGGAGCCACGCGCCAGAGAGCTGGTGATGATGACTGATGATGATGATGATG 165  
Oy 42 GlyProCysArgPheMetAlaProValLeuAlaGluIleAlaLysLysThrProGluLeu 61  
Db 166 GGACCATGCGCCATGATGCTGCTCAATTTTCTGATGCTGCGCAGAAAGTTCACGCTGCT 225  
Oy 62 IlePheLeuLysValAspValAspGluValArgProValAlaGluGluTrpSerIleGlu 81  
Db 226 GTTTTCTCAAGGTCGAGCTGATGATGATGATGATGATGATGATGATGATGATGATG 285  
Oy 82 AlaMetProThrPheLeuPheLeuLysAspGluIleValAspLysValAlaGlyVal 101  
Db 286 GCCATGCCACACCTTCCTGTTTCATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 345  
Oy 102 SerLysAspAspLeuGlnAlaThrIleAlaLysAlaSerAla 116  
Db 346 ATCAAGGAGAGAGCTGACAGCAGGATGCTGCTCCAGCGGCTGCC 390

RESULT 2  
US-08-181-271A-103

Sequence 103, Application US/08181271A

Patent No. 5614395

GENERAL INFORMATION:

APPLICANT: Ryals, John A.  
APPLICANT: Alexander, Danny C.  
APPLICANT: Beck, James J.  
APPLICANT: Duesing, John H.  
APPLICANT: Friedrich, Leslie B.  
APPLICANT: Goodman, Robert M.  
APPLICANT: Harms, Christian  
APPLICANT: Melns, Jr., Frederick  
APPLICANT: Montoya, Alice  
APPLICANT: Moyer, Mary B.  
APPLICANT: Neuhaus, Jean-Marc  
APPLICANT: Payne, George B.  
APPLICANT: Spertison, Christoph  
APPLICANT: Stinson, Jeffrey R.  
APPLICANT: Uknes, Scott J.  
APPLICANT: Ward, Eric R.  
APPLICANT: Williams, Sherica C.  
TITLE OF INVENTION: CHEMICALLY REGULATABLE AND ANTI-PATHOGENIC  
NUMBER OF SEQUENCES: 106  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: CIBA-GEIGY Corporation  
STREET: 7 Skyline Drive  
CITY: Hawthorne  
STATE: New York  
COUNTRY: USA  
ZIP: 10532

COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/181,271A  
FILING DATE: 13-JAN-94  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 08/093,301  
FILING DATE: 16-JUL-1993  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/937,197  
FILING DATE: 6-NOV-1992  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/678,378  
FILING DATE: 1-APR-1991  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/305,566  
FILING DATE: 6-FEB-1989  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/165,667  
FILING DATE: 8-MAR-1988  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/042,847  
FILING DATE: 6-APR-1993  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/632,441  
FILING DATE: 21-DEC-1990  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/425,504  
FILING DATE: 20-OCT-1989  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/848,506  
FILING DATE: 6-MAR-1992  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/768,122  
FILING DATE: 27-SEP-1991  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/580,431  
FILING DATE: 7-SEP-1990  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/368,672  
FILING DATE: 20-JUN-1989  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/045,957  
FILING DATE: 12-APR-1993  
ATTORNEY/AGENT INFORMATION:  
NAME: Elmer, James Scott  
REGISTRATION NUMBER: 36,129  
REFERENCE/DOCKET NUMBER: S-19825/P1/CGC 1727  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (919)541-8614  
TELEFAX: (919)541-8689  
INFORMATION FOR SEQ ID NO: 103:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 653 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: DNA (genomic)  
US-08-181-271A-103

Alignment Scores:

Pred. No.: 1,29e-45 Length: 653  
Score: 392.00 Matches: 72  
Percent Similarity: 83.04% Conservative: 21  
Best Local Similarity: 64.29% Mismatches: 18  
Query Match: 63.43% Indels: 1  
DB: 1 Gaps: 0

US-09-786-715-8 (1-123) x US-08-181-271A-103 (1-653)

Oy 5 GluGluGluGluValIleGlyValHisThrValAspGluTrpLeuGlnLeuGlnAsn 24  
Db 75 GAGGAGGACAAAGTTCGCTGCGCACAGGTTGAGGATGAGAGAGAGAGAGAGAGAG 134  
Oy 25 AlaLysAspSerLysLeuIleValValAspPheThrAlaSerTrpCysGlyProCys 44

Db 135 GGGCTTGAGACTAAGAACTGCTGCTGCTGCTTACTTCTTCATGCTGCGGSCCTGC 194  
QY 45 ArgpHemetaIaprovaleuAlagluileAlaIySthrProgluIeuIlephelu 64  
Db 195 CGTTTATGCCCCCAATTCCTGCTGACATTGCTAAGAAAGTGCATGTAATTCCTC 254  
QY 65 LysValaspValaspIuValargProvalAla-GluIuTySerIlegluAlaMetPr 84  
Db 255 AAGCTTGATGCTGATGACTGAGAGACTGTTTCAGCGGGAATGAGTGTGAGGCATGCC 314  
QY 84 orthrPheuleuPheuleuLysaspIygluIleValaspLysValalglYAlaSerLysas 104  
Db 315 AACTTTGCTTCATTAAGATGAAGAAAGAGTGAAGAGTGTGTCGCAAGAAAGA 374  
QY 104 paspleuGlAlaThrIleAlaIyShsAlaSer 115  
Db 375 GGAGTTGCAGACACCATAGTGAAGCATGCTGCT 408

RESULT 3  
US-08-449-315-103  
Sequence 103, Application US/08449315  
Patent No. 5650305  
GENERAL INFORMATION:  
APPLICANT: Ryals, John A.  
APPLICANT: Alexander, Danny C.  
APPLICANT: Beck, James J.  
APPLICANT: Duesing, John H.  
APPLICANT: Friedrich, Leslie B.  
APPLICANT: Goodman, Robert M.  
APPLICANT: Harms, Christian  
APPLICANT: Meins, Jr., Frederick  
APPLICANT: Montoya, Alice  
APPLICANT: Moyer, Mary B.  
APPLICANT: Neuhaus, Jean-Marc  
APPLICANT: Payne, George B.  
APPLICANT: Sperison, Christoph  
APPLICANT: Stinson, Jeffrey R.  
APPLICANT: Uknes, Scott J.  
APPLICANT: Ward, Eric R.  
APPLICANT: Williams, Shericea C.  
TITLE OF INVENTION: CHEMICALLY REGULATABLE AND ANTI-PATHOGENIC  
NUMBER OF SEQUENCES: 106  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: CIBA-GEIGY Corporation  
STREET: 7 Skyline Drive  
CITY: Hawthorne  
STATE: New York  
COUNTRY: USA  
ZIP: 10532  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/449,315  
FILING DATE: 24-MAY-1995  
CLASSIFICATION: 800  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/181,271  
FILING DATE: 13-JAN-94  
APPLICATION NUMBER: US 08/093,301  
FILING DATE: 16-JUL-1993  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/937,197  
FILING DATE: 6-NOV-1992  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/678,378  
FILING DATE: 1-APR-1991  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/305,566  
FILING DATE: 6-FEB-1989

PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/165,667  
FILING DATE: 8-MAR-1988  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/042,847  
FILING DATE: 6-APR-1993  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/632,441  
FILING DATE: 21-DEC-1990  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/425,504  
FILING DATE: 20-OCT-1989  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/848,506  
FILING DATE: 6-MAR-1992  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/768,122  
FILING DATE: 27-SEP-1991  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/580,431  
FILING DATE: 7-SEP-1990  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/368,672  
FILING DATE: 20-JUN-1989  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/329,018  
FILING DATE: 24-MAR-1989  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/045,957  
FILING DATE: 12-APR-1993  
ATTORNEY/AGENT INFORMATION:  
NAME: Elmer, James Scott  
REGISTRATION NUMBER: 36,129  
REFERENCE/DOCKET NUMBER: S-19825/P1/CGC 1727  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (919)541-8614  
TELEFAX: (919)541-8689  
INFORMATION FOR SEQ ID NO: 103:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 653 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: DNA (genomic)  
US-08-449-315-103

Alignment Scores:  
Pred. No.: 1,296-45  
Score: 392.00  
Percent Similarity: 83.048  
Best Local Similarity: 64.298  
Query Match: 63.434  
DB: 1  
Gaps: 0

US-09-786-715-8 (1-123) x US-08-449-315-103 (1-653)

QY 5 GluIuGlYgluValIleglYValHIsThrValaspIuTrpLysleuIuengIuAsn 24  
Db 75 GAGGAGGAGCAAGTTCGCTGCCACAGCTTGAGAGTGAACGACTTACTTACAGAA 134  
QY 25 AlAlYsAspSerLysleuIleValaspPheThrAlaSerTrpGlyProGAs 44  
Db 135 GGGCTTGAGACTAAGAACTGCTGCTGCTGCTTACTTCTTCATGCTGCGGSCCTGC 194  
QY 45 ArgpHemetaIaprovaleuAlagluileAlaIySthrProgluIeuIlephelu 64  
Db 195 CGTTTATGCCCCCAATTCCTGCTGACATTGCTAAGAAAGTGCATGTAATTCCTC 254  
QY 65 LysValaspValaspIuValargProvalAla-GluIuTySerIlegluAlaMetPr 84  
Db 255 AAGCTTGATGCTGATGACTGAGAGACTGTTTCAGCGGGAATGAGTGTGAGGCATGCC 314  
QY 84 orthrPheuleuPheuleuLysaspIygluIleValaspLysValalglYAlaSerLysas 104







```

1      CITY: Tarrytown
2      STATE: New York
3      COUNTRY: USA
4      ZIP: 10591
5
6      COMPUTER READABLE FORM:
7      MEDIUM TYPE: Floppy disk
8      COMPUTER: IBM PC compatible
9      OPERATING SYSTEM: PC-DOS/MS-DOS
10     SOFTWARE: Patent in Release #1.0, Version #1.25
11     CURRENT APPLICATION DATA:
12     APPLICATION NUMBER: US/08/456,265A
13     FILING DATE: 31-MAY-95
14     CLASSIFICATION: 435
15     PRIOR APPLICATION DATA:
16     APPLICATION NUMBER: US 08/181,271
17     FILING DATE: 13-JAN-1994
18     PRIOR APPLICATION DATA:
19     APPLICATION NUMBER: US 08/093,301
20     FILING DATE: 16-JUL-1993
21     PRIOR APPLICATION DATA:
22     APPLICATION NUMBER: US 07/937,197
23     FILING DATE: 6-NOV-1992
24     PRIOR APPLICATION DATA:
25     APPLICATION NUMBER: US 07/678,378
26     FILING DATE: 1-APR-1991
27     PRIOR APPLICATION DATA:
28     APPLICATION NUMBER: US 07/305,566
29     FILING DATE: 6-FEB-1989
30     PRIOR APPLICATION DATA:
31     APPLICATION NUMBER: US 07/165,667
32     FILING DATE: 8-MAR-1988
33     PRIOR APPLICATION DATA:
34     APPLICATION NUMBER: US 08/042,847
35     FILING DATE: 6-APR-1993
36     PRIOR APPLICATION DATA:
37     APPLICATION NUMBER: US 07/632,441
38     FILING DATE: 21-DEC-1990
39     PRIOR APPLICATION DATA:
40     APPLICATION NUMBER: US 07/425,504
41     FILING DATE: 20-OCT-1989
42     PRIOR APPLICATION DATA:
43     APPLICATION NUMBER: US 07/848,506
44     FILING DATE: 6-MAR-1992
45     PRIOR APPLICATION DATA:
46     APPLICATION NUMBER: US 07/768,122
47     FILING DATE: 27-SEP-1991
48     PRIOR APPLICATION DATA:
49     APPLICATION NUMBER: US 07/580,431
50     FILING DATE: 7-SEP-1990
51     PRIOR APPLICATION DATA:
52     APPLICATION NUMBER: US 07/368,672
53     FILING DATE: 20-JUN-1989
54     PRIOR APPLICATION DATA:
55     APPLICATION NUMBER: US 07/329,018
56     FILING DATE: 24-MAR-1989
57     PRIOR APPLICATION DATA:
58     APPLICATION NUMBER: US 08/045,957
59     FILING DATE: 12-APR-1993
60     ATTORNEY/AGENT INFORMATION:
61     NAME: Meligs, J. Timothy
62     REGISTRATION NUMBER: 38, 241
63     REFERENCE/DOCKET NUMBER: S-19825/P1/CGC 1727/DIV10
64     TELECOMMUNICATION INFORMATION:
65     TELEPHONE: (919)541-8587
66     TELEFAX: (919)541-8669
67     INFORMATION FOR SEQ ID NO: 103:
68     SEQUENCE CHARACTERISTICS:
69     LENGTH: 653 base pairs
70     TYPE: nucleic acid
71     STRANDEDNESS: single
72     TOPOLOGY: linear
73     MOLECULE TYPE: DNA (genomic)
74     OS-08-456-265A-103

```

APPLICATION NUMBER: US/08/455,416  
FILING DATE: 31-MAY-1995  
CLASSIFICATION: 800  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/181,271  
FILING DATE: 13-JAN-94  
APPLICATION NUMBER: US 08/093,301  
FILING DATE: 16-JUL-1993  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/937,197  
FILING DATE: 6-NOV-1992  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/678,378  
FILING DATE: 1-APR-1991  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/305,566  
FILING DATE: 6-FEB-1989  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/165,667  
FILING DATE: 8-MAR-1988  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/042,847  
FILING DATE: 6-APR-1993  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/632,441  
FILING DATE: 21-DEC-1990  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/425,504  
FILING DATE: 20-OCT-1989  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/848,506  
FILING DATE: 6-MAR-1992  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/768,122  
FILING DATE: 27-SEP-1991  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/580,431  
FILING DATE: 7-SEP-1990  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/368,672  
FILING DATE: 20-JUN-1989  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/329,018  
FILING DATE: 24-MAR-1989  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/045,957  
FILING DATE: 12-APR-1993  
ATTORNEY/AGENT INFORMATION:  
NAME: Elmer, James Scott  
REGISTRATION NUMBER: 36,129  
REFERENCE/DOCKET NUMBER: S-19825/P1/CGC 1727  
TELEPHONE: (919)541-8614  
TELEFAX: (919)541-8689  
INFORMATION FOR SEQ ID NO: 103:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 653 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: DNA (genomic)  
US-08-455-416-103

Alignment Scores:  
Pred. No.: 1.29e-45 Length: 653  
Score: 392.00 Matches: 72  
Percent Similarity: 83.04% Conservative: 21  
Best Local Similarity: 64.29% Mismatches: 18  
Query Match: 63.43% Indels: 1  
DB: 1 Gaps: 0

US-09-786-715-8 (1-123) x US-08-455-416-103 (1-653)

QY 5 GluGluGlyGlnValIleGlyValHisThrValAspGluTrpLysLeuGlnLeuGlnAsn 24  
|||||  
Db 75 GAGGAGGACCAAGTTCGGCTCCACCAAGCTTGAGCAAGCAAGCACTTCAAGAAA 134  
QY 25 AlAlAspSerLysLysLeuIleValAspPheThrAlaSerTrpCysGlyProCys 44  
:|||||  
Db 135 GCGGTGAGACTAGAAAGTGGGTGGTGGATTTTACGCTTCATGCTGCGGSCCTTGC 194  
QY 45 ArgPheMetAlaProValLeuAlaGluIleAlaLysThrProGluLeuIlePheLeu 64  
|||||  
Db 195 CGTTTTATTGCCCATTCCTGCACTTAAGAAAGATGCCCATGTTATATTCCTC 254  
QY 65 LysValAspValAspGluValArgProValAla-GluGluTrpSerIleGluAlaMetPr 84  
|||||  
Db 255 AAGGTGATGTTGATGACGAGTGTTCACCGGAGATGAGTGTGGAGGCAATGCC 314  
QY 84 oThrPheLeuPheLeuLysAspLysLysLeuIleValAspLysValAlaGlyAlaSerLysAs 104  
|||||  
Db 315 AACTTTGTCTTCATTAAAGATGGAAGAAAGAGGACAGAGTGTGTGTCACAAAGAA 374  
QY 104 pasPleuGlnAlaThrIleAlaLysHisAlaSer 115  
:|||||  
Db 375 GGAGTTGCAGACAGCACCATAGTGAACATCTCTGCT 408

RESULT 8  
US-08-455-244-103  
Sequence 103, Application US/08455244  
Patent No. 5789214  
GENERAL INFORMATION:  
APPLICANT: Ryals, John A.  
APPLICANT: Alexander, Danny C.  
APPLICANT: Beck, James J.  
APPLICANT: Duesting, John H.  
APPLICANT: Friedrich, Leslie B.  
APPLICANT: Goodman, Robert M.  
APPLICANT: Harms, Christian  
APPLICANT: Meins, Jr., Frederick  
APPLICANT: Montoya, Alice  
APPLICANT: Moyer, Mary B.  
APPLICANT: Neuhaus, Jean-Marc  
APPLICANT: Payne, George B.  
APPLICANT: Sperison, Christoph  
APPLICANT: Stinson, Jeffrey R.  
APPLICANT: Uknes, Scott J.  
APPLICANT: Ward, Eric R.  
TITLE OF INVENTION: CHEMICALLY REGULATABLE AND ANTI-PATHOGENIC  
TITLE OF INVENTION: DNA SEQUENCES AND USES THEREOF  
NUMBER OF SEQUENCES: 106  
CORRESPONDENCE ADDRESS:  
ADDRESSER: CIBA-GEIGY Corporation  
STREET: 7 Skyline Drive  
CITY: Hawthorne  
STATE: New York  
COUNTRY: USA  
ZIP: 10532  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/455,244  
FILING DATE: 31-MAY-1995  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/181,271  
FILING DATE: 13-JAN-94  
APPLICATION NUMBER: US 08/093,301  
FILING DATE: 16-JUL-1993  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/937,197  
FILING DATE: 6-NOV-1992

PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/678,378  
FILING DATE: 1-APR-1991  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/305,566  
FILING DATE: 6-FEB-1989  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/165,667  
FILING DATE: 8-MAR-1988  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/042,847  
FILING DATE: 6-APR-1993  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/632,441  
FILING DATE: 21-DEC-1990  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/425,504  
FILING DATE: 20-OCT-1989  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/848,506  
FILING DATE: 6-MAR-1992  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/768,122  
FILING DATE: 27-SEP-1991  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/580,431  
FILING DATE: 7-SEP-1990  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/368,672  
FILING DATE: 20-JUN-1989  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/329,018  
FILING DATE: 24-MAR-1989  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/045,957  
FILING DATE: 12-APR-1993  
ATTORNEY/AGENT INFORMATION:  
NAME: Elmer, James Scott  
REGISTRATION NUMBER: 36,129  
REFERENCE/DOCKET NUMBER: S-19825/P1/CGC 1727  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (919)541-8614  
TELEFAX: (919)541-8689  
INFORMATION FOR SEQ ID NO: 103:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 653 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: DNA (genomic)  
US-08-455-244-103

Alignment Scores:  
Pred. No.: 1 296-45 Length: 653  
Score: 392.00 Matches: 72  
Percent Similarity: 83.04% Conservative: 21  
Best Local Similarity: 64.29% Mismatches: 18  
Query Match: 63.43% Indels: 1  
DB: 1 Gaps: 0

US-09-786-715-8 (1-123) x US-08-455-244-103 (1-653)

QY 5 GlnGlnGlnGlnValIleGlnValAlaIsthrValAspGluTrpLysLeuGlnAsn 24  
Db 75 GAGAGGAGCAAGAGTTCGGCTGCACAGAGTTCAGAGTTCAGAGTTCAGAGAA 134  
QY 25 AlAlAspSerLysLysLeuIleValAlaAspPheThrAlaSerTrpCysGlyProCys 44  
Db 135 GCGCTTGACGACTAGAAAGTGGTGGTGCATTTTACTGCTCATGGTGGGSCCTTGC 194  
QY 45 ArpPheMetIaIaProValLeuAlaGlnIleAlaLysLysTrpProGlnLeuIlePheLeu 64  
Db 195 CGTTTATGCGCCCAATTCCTTGTGACATTCGTAAAGAGAGCCCATGTTATATCTCC 254

QY 65 LysValAspValAspGluValArpProValAla-GlnGlnTrpSerIleGlnAlaMetPr 84  
Db 255 AAGCTTGATGTTGATGAAGACTGAGAGCTTTTCAGCGGGAATGAGATGTGAGCGCAATGCC 314  
QY 84 OthrPheLeuPheLeuLysAspGlyGlnIleValAlaAspLysValAlaGlnAlaSerLysAs 104  
Db 315 AACTTTTCTCTCATTAAGATGGAAGAAGATGGAGAGAGTTGTTGGTGCACAGAAAGA 374  
QY 104 PAspLeuGlnAlaThrIleAlaLysHisAlaSer 115  
Db 375 GGAGTTGACGACGACCATAGTGAAGCATGCTGCT 408

RESULT 9  
US-08-454-876-103  
Sequence 103 Application US/08454876  
Patent No. 3804693  
GENERAL INFORMATION:  
APPLICANT: Ryals, John A.  
APPLICANT: Alexander, Danny C.  
APPLICANT: Beck, James J.  
APPLICANT: Duesing, John H.  
APPLICANT: Friedrich, Leslie B.  
APPLICANT: Goodman, Robert M.  
APPLICANT: Harms, Christian  
APPLICANT: Meins, Jr., Frederick  
APPLICANT: Montoya, Alice  
APPLICANT: Moyer, Mary B.  
APPLICANT: Neuhaus, Jean-Marc  
APPLICANT: Payne, George B.  
APPLICANT: Sperison, Christoph  
APPLICANT: Stinson, Jeffrey R.  
APPLICANT: Uknes, Scott J.  
APPLICANT: Ward, Eric R.  
APPLICANT: Williams, Shericca C.  
TITLE OF INVENTION: CHEMICALLY REGULATABLE AND ANT1-PATHOGENIC  
TITLE OF INVENTION: DNA SEQUENCES AND USES THEREOF  
NUMBER OF SEQUENCES: 106  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: CIBA-GEIGY Corporation  
STREET: 7 Skyline Drive  
CITY: Hawthorne  
STATE: New York  
COUNTRY: USA  
ZIP: 10532  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/454,876  
FILING DATE: 31-MAY-1995  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/181,271  
FILING DATE: 13-JAN-94  
APPLICATION NUMBER: US 08/093,301  
FILING DATE: 16-JUL-1993  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/937,197  
FILING DATE: 6-NOV-1992  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/678,378  
FILING DATE: 1-APR-1991  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/305,566  
FILING DATE: 6-FEB-1989  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/165,667  
FILING DATE: 8-MAR-1988  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/042,847



APPLICATION NUMBER: US 07/768,122  
FILING DATE: 27-SEP-1991  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/580,431  
FILING DATE: 7-SEP-1990  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/368,672  
FILING DATE: 20-JUN-1989  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/329,018  
FILING DATE: 24-MAR-1989  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/045,957  
FILING DATE: 12-APR-1993  
ATTORNEY/AGENT INFORMATION:  
NAME: Elmer, James Scott  
REGISTRATION NUMBER: 36,129  
REFERENCE/DOCKET NUMBER: S-19825/P1/CGC 1727  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (919)541-8614  
TELEFAX: (919)541-8689  
INFORMATION FOR SEQ ID NO: 103:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 653 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: DNA (genomic)  
US-08-457-364-103

Alignment Scores:  
Pred. No.: 1.29e-45 Length: 653  
Score: 392.00 Matches: 72  
Percent Similarity: 83.04% Conservative: 21  
Best Local Similarity: 64.29% Mismatches: 18  
Query Match: 63.43% Indels: 1  
DB: 2 Gaps: 0

US-09-786-715-8 (1-123) x US-08-457-364-103 (1-653)

QY 5 GluGlulGlyGlnValIleGlyValHisThrValAspGluTrpLysLeuGlnLeuGlnasn 24  
|||||  
DB 75 GAGGAGGACACAGCTTCGGCTGCCACAGGTTGAGCAATGACGAGTACTTCAAGAA 134  
|||||  
QY 25 AAlaLysAspSerLysLysLeuIleValValAspSerThrProGluLeuIlePheleu 44  
|||||  
DB 135 GGCCTGAGACTAGAAACGTCGTGTCGATTTTCTCATGTCGTCGCGGSCCTTGC 194  
|||||  
QY 45 ArgPheMetLarProValIleAlaGluIleAlaLysLysThrProGluLeuIlePheleu 64  
|||||  
DB 195 CGTTTATTTGGCCCAATTCCTGTCGACATTCCTAAGAAATGCCCATGTTATTTCTTC 254  
|||||  
QY 65 LysValAspValAspGluValArgProValAla-GluGluTrpSerIleGluAlaMetPr 84  
|||||  
DB 255 AAGGTTATGTTGATGAACTGAACACTGTTTCACCGGAATGAGAGTGTGGAGCAATGCC 314  
|||||  
QY 84 othPheLeuPheLeuLysAspGlyGluIleValAspLysValValGlyAlaSerLysas 104  
|||||  
DB 315 AACCTTTGCTCTTCATTAAGATGGAAGAAAGATGAGAGAGTGTGTGTCAGAAAGA 374  
|||||  
QY 104 pAspLeuGlnAlaThrIleAlaLysHisAlaSer 115  
|||||  
DB 375 GGAGTTCAGCAGCACCATAGTGAAGCATGCTGCT 408  
|||||

RESULT 11  
US-08-456-262-103  
Sequence 103. Application US/08456262  
Patent No. 5851766  
GENERAL INFORMATION:  
APPLICANT: Ryals, John A.  
APPLICANT: Alexander, Danny C.  
APPLICANT: Beck, James J.  
APPLICANT: Duesing, John H.

APPLICANT: Friedrich, Leslie B.  
APPLICANT: Goodman, Robert M.  
APPLICANT: Harms, Christian  
APPLICANT: Meins, Jr., Frederick  
APPLICANT: Montoya, Alice  
APPLICANT: Moyer, Mary B.  
APPLICANT: Neunhaus, Jean-Marc  
APPLICANT: Payne, George B.  
APPLICANT: Spertson, Christoph  
APPLICANT: Stinson, Jeffrey R.  
APPLICANT: Uknes, Scott J.  
APPLICANT: Ward, Eric R.  
APPLICANT: Williams, Shericca C.  
TITLE OF INVENTION: CHEMICALLY REGULATABLE AND ANTI-PATHOGENIC  
TITLE OF INVENTION: DNA SEQUENCES AND USES THEREOF  
NUMBER OF SEQUENCES: 106  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: CIBA-GEIGY Corporation  
STREET: 7 Skyline Drive  
CITY: Hawthorne  
STATE: New York  
COUNTRY: USA  
ZIP: 10532  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/456,262  
FILING DATE: 31-MAY-1995  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/181,271  
FILING DATE: 13-JAN-94  
APPLICATION NUMBER: US 08/093,301  
FILING DATE: 16-JUL-1993  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/937,197  
FILING DATE: 6-NOV-1992  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/678,378  
FILING DATE: 1-APR-1991  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/305,566  
FILING DATE: 6-FEB-1989  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/165,667  
FILING DATE: 8-MAR-1988  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/042,847  
FILING DATE: 6-APR-1993  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/632,441  
FILING DATE: 21-DEC-1990  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/425,504  
FILING DATE: 20-OCT 1989  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/848,506  
FILING DATE: 6-MAR-1992  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/768,122  
FILING DATE: 27-SEP-1991  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/580,431  
FILING DATE: 7-SEP-1990  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/368,672  
FILING DATE: 20-JUN-1989  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/329,018  
FILING DATE: 24-MAR-1989

PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/045,957  
FILING DATE: 12-APR-1993  
ATTORNEY/AGENT INFORMATION:  
NAME: Elmer, James Scott  
REGISTRATION NUMBER: 36,129  
REFERENCE/DOCKET NUMBER: S-19825/P1/CGC 1727  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (919)541-8614  
TELEFAX: (919)541-8689  
INFORMATION FOR SEQ ID NO: 103:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 653 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: DNA (genomic)  
US-08-456-262-103

Alignment Scores:  
Pred. No.: 1,296-45 Length: 653  
Score: 392.00 Matches: 72  
Percent Similarity: 83.04% Conservative: 21  
Best Local Similarity: 64.29% Mismatches: 18  
Query Match: 63.43% Indels: 1  
DB: 2 Gaps: 0

US-09-786-715-8 (1-123) x US-08-456-262-103 (1-653)

QY 5 GIUGIUGIYNVALIIEGLIYVAIHSTHVALASPGIUTRPLYLEUGINLEUGINAsn 24  
|||||  
DB 75 GAGGAGGAGCAGTGTGGCGCCACAGGTGAGAGTGAACAGACTTCAAGAA 134  
|||||

QY 25 AIALYASPSerLYSLLEULEVALVALASpHEThALSetRTPcysGLYProCys 44  
|||||

DB 135 GCGGTGAGACTAAGAACTGCTGCTGCTGCTTACTGCTTCAATGCTGCGGSCCTTGC 194  
|||||

QY 45 ARGpHEMELAPROVALLEULAGIULEALALYSLYSTHProGLUleuLEPHELEu 64  
|||||

DB 195 CGTTTATTTGCCCATTTCTTCTGACATGCTAAGAAAGATCCCATGTTATTTCCCTC 254  
|||||

QY 65 LYSVALASpVALASpGLIUVAlARpPROVALAlA-GLIULYrSerILEGLIAlAMePr 84  
|||||

DB 255 AAGGTGAGTGTGATGAGACTGAAGACTGTTTACGCGGGAATGAGTGTGAGAGCAATGCC 314  
|||||

QY 84 OThrPHEuPHEuLYSLYSLYGLIULEVALASpLYSVALVALGLYAlASerLYSAs 104  
|||||

DB 315 AACTTTGCTCTCATTTAAAGATGAAAGAGTGAAGAGAGTTGTTGTCACAGAAAGA 374  
|||||

QY 104 PASpLEUGINAlATHrILLeALALYSHISLASer 115  
|||||

DB 375 GGAGTTGCAGCAGACCATAGTAGAGCATGCTGCT 408  
|||||

RESULT 12  
US-08-456-240-103  
Sequence 103, Application US/08456240  
Patient No. 5856154  
GENERAL INFORMATION:  
APPLICANT: Ryals, John A.  
APPLICANT: Alexander, Danny C.  
APPLICANT: Beck, James J.  
APPLICANT: Duesing, John H.  
APPLICANT: Friedrich, Leslie B.  
APPLICANT: Goodman, Robert M.  
APPLICANT: Harms, Christian  
APPLICANT: Meins, Jr., Frederick  
APPLICANT: Montoya, Alice  
APPLICANT: Moyer, Mary B.  
APPLICANT: Neuhaus, Jean-Marc  
APPLICANT: Payne, George B.  
APPLICANT: Sperison, Christoph  
APPLICANT: Stinson, Jeffrey R.  
APPLICANT: Uknes, Scott J.

APPLICANT: Ward, Eric R.  
APPLICANT: Williams, Shericka C.  
TITLE OF INVENTION: CHEMICALLY REGULATABLE AND ANTI-PATHOGENIC  
TITLE OF INVENTION: DNA SEQUENCES AND USES THEREOF  
NUMBER OF SEQUENCES: 106  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: CIBA-GEIGY Corporation  
STREET: 7 Skyline Drive  
CITY: Hawthorne  
STATE: New York  
COUNTRY: USA  
ZIP: 10532

COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/456,240  
FILING DATE: 31-MAY-1995

CLASSIFICATION: 800  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/181,271  
FILING DATE: 13-JAN-94  
APPLICATION NUMBER: US 08/093,301  
FILING DATE: 16-JUL-1993

PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/937,197  
FILING DATE: 6-NOV-1992

PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/678,378  
FILING DATE: 1-APR-1991

PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/305,566  
FILING DATE: 6-FEB-1989

PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/165,667  
FILING DATE: 8-MAR-1988

PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/042,847  
FILING DATE: 6-APR-1993

PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/632,441  
FILING DATE: 21-DEC-1990

PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/425,504  
FILING DATE: 20-OCT 1989

PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/848,506  
FILING DATE: 6-MAR-1992

PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/768,122  
FILING DATE: 27-SEP-1991

PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/580,431  
FILING DATE: 7-SEP-1990

PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/368,672  
FILING DATE: 20-JUN-1989

PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/329,018  
FILING DATE: 24-MAR-1989

PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/045,957  
FILING DATE: 12-APR-1993

ATTORNEY/AGENT INFORMATION:  
NAME: Elmer, James Scott  
REGISTRATION NUMBER: 36,129  
REFERENCE/DOCKET NUMBER: S-19825/P1/CGC 1727  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (919)541-8614  
TELEFAX: (919)541-8689  
INFORMATION FOR SEQ ID NO: 103:

SEQUENCE CHARACTERISTICS:  
LENGTH: 653 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: DNA (genomic)  
US-08-456-240-103

Alignment Scores:  
Pred. No.: 1,29e-45 Length: 653  
Score: 392.00 Matches: 72  
Percent Similarity: 83.04% Conservative: 21  
Best Local Similarity: 64.29% Mismatches: 18  
Query Match: 63.43% Indels: 1  
Gaps: 0

US-09-786-715-8 (1-123) x US-08-456-240-103 (1-653)

QY 5 GluGluGlyGlnValIleGlyValHisThrValAspGluTrpLysLeuGlnLeuGlnAsn 24  
Db 75 GAGGAGGAGCAAGAGTTCGGCTGCCACAGGTTGAGAGATGCAAGCAGTACTTCAAGAAA 134  
QY 25 AlalysAspSerLysLysLeuIleValValAspPheThrAlaSerTrpCysGlyProCys 44  
Db 135 GGCCTTACAGACTAGAACTGGTGGTGCATTTTACTGCTGATGATGGGSCCTTGC 194  
QY 45 ArgPheMetAlaProValLeuAlaGluIleAlaLysLysThrProGluLeuIlePheLeu 64  
Db 195 CGTTTATTCGCCCAATCTTCTGCATCTGCTAAGAGATGCCCATGTTATATTTCTC 254  
QY 65 LysValAspValAspGluValArgProValAla-GluGluTrpSerIleGluAlaMetPr 84  
Db 255 AAGGTGATGTGATGATGAGAGACTGTTTCACGGGAAATGAGGTGGAGCAATGCC 314  
QY 84 othrPheLeuPheLeuLysAspGlyGluIleValAspLysValValGlyAlaSerLysAs 104  
Db 315 AACTTTGTCTCTTAAAGATGGAAGAAGTGAAGATGTTGTGTGCCAAGAAA 374  
QY 104 PaspLeuGlnAlaThrIleAlaLysHisAlaSer 115  
Db 375 GGAATTCCAGACACCATAGTGAAGCATGCTGCT 408

RESULT 13  
US-08-455-736-103  
Sequence 103, Application US/08455736  
Patent No. 5880328  
GENERAL INFORMATION:  
APPLICANT: Ryals, John A.  
APPLICANT: Alexander, Danny C.  
APPLICANT: Beck, James J.  
APPLICANT: Duesing, John H.  
APPLICANT: Friedrich, Leslie B.  
APPLICANT: Goodman, Robert M.  
APPLICANT: Harms, Christian  
APPLICANT: Melns, Jr., Frederick  
APPLICANT: Montoya, Alice  
APPLICANT: Moyer, Mary B.  
APPLICANT: Neuhaus, Jean-Marc  
APPLICANT: Payne, George B.  
APPLICANT: Sperison, Christoph  
APPLICANT: Stinson, Jeffrey R.  
APPLICANT: Uknes, Scott J.  
APPLICANT: Ward, Eric R.  
APPLICANT: Williams, Shericea C.  
TITLE OF INVENTION: CHEMICALLY REGULATABLE AND ANTI-PATHOGENIC  
TITLE OF INVENTION: DNA SEQUENCES AND USES THEREOF  
NUMBER OF SEQUENCES: 106  
CORRESPONDENCE ADDRESSES:  
ADDRESSEE: CIBA-GEIGY Corporation  
STREET: 7 Skyline Drive  
CITY: Hawthorne  
STATE: New York  
COUNTRY: USA

ZIP: 10532  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/455,736  
FILING DATE: 31-MAY-1995  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/181,271  
FILING DATE: 13-JAN-1994  
APPLICATION NUMBER: US 08/093,301  
FILING DATE: 16-JUL-1993  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/937,197  
FILING DATE: 6-NOV-1992  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/678,378  
FILING DATE: 1-APR-1991  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/305,566  
FILING DATE: 6-FEB-1989  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/165,667  
FILING DATE: 8-MAR-1988  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/042,847  
FILING DATE: 6-APR-1993  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/632,441  
FILING DATE: 21-DEC-1990  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/425,504  
FILING DATE: 20-OCT-1989  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/848,506  
FILING DATE: 6-MAR-1992  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/768,122  
FILING DATE: 27-SEP-1991  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/580,431  
FILING DATE: 7-SEP-1990  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/368,672  
FILING DATE: 20-JUN-1989  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/329,018  
FILING DATE: 24-MAR-1989  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/045,957  
FILING DATE: 12-APR-1993  
ATTORNEY/AGENT INFORMATION:  
NAME: Elmer, James Scott  
REGISTRATION NUMBER: 36,129  
REFERENCE/DOCKET NUMBER: S-19825/P1/CGC 1727  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (919)541-8614  
TELEFAX: (919)541-8689  
INFORMATION FOR SEQ ID NO: 103:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 653 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: DNA (genomic)  
US-08-455-736-103

Alignment Scores:  
Pred. No.: 1,29e-45 Length: 653  
Score: 392.00 Matches: 72



Percent Similarity: 83.04% Conservative: 21  
Best Local Similarity: 64.29% Mismatches: 18  
Query Match: 63.43% Indels: 1  
DB: 2 Gaps: 0

US-09-786-715-8 (1-123) x US-08-455-736-103 (1-653)

OY 5 GUGUGUGLGNValIILeGlyValHsThrValAspGluTrpLysLeuGlnLeuGlnasn 24  
|||||  
DB 75 GAGGAGGACACAGTGTTCGCTGCCACACAGTTGGAGATGGACGAGTACTTCAAGAAA 134  
OY 25 AlalysAspSerLysLysLeuIleValAlAspPheThrAlaSerTrpCysGlyProCys 44  
|||||  
DB 135 GCCGTTGAGACTAAAGAACTGGTGGTGCATTTTACTGCTCATGTGGCGGSCCTTGC 194  
OY 45 ArgPheMetAlaProValLeuAlaGluIleAlaLysLysThrProGluLeuIlePheLeu 64  
|||||  
DB 195 GCTTTATTGCCCCCAATTCCTGCTGACATTCCTAAGAAAGATGCCCATGTATTATTCCTC 254  
OY 65 LysValAspValAspGluValArgProValAla-GluGluTrpSerIleGluAlaMetPr 84  
|||||  
DB 255 AAGTTGATGTTGATGAACACTGAGACTGTTTCAGCGGGAATGAGTGTGGAGCAATGCC 314  
OY 84 OThrPheLeuPheLeuLysAspGlyGluIleValAspLysValAlaGlyAlaSerLysAs 104  
|||||  
DB 315 AACTTTGTCTTCATTAAAGATGGAAGAAAGTGCACAGAGTTGTTGTTGCCAAGAAAAGA 374  
OY 104 PASPheGlnAlaThrIleAlaLysHisAlaSer 115  
|||||  
DB 375 GGAGTTGCAGCAGACCATAGTGAAGCATGCTGCT 408

RESULT 14  
US-08-971-217-103  
Sequence 103 Application US/08971217  
Patent No. 5942662  
GENERAL INFORMATION:  
APPLICANT: Ryals, John A.  
APPLICANT: Harms, Christian  
APPLICANT: Friedrich, Leslie  
APPLICANT: Beck, James  
APPLICANT: Uknes, Scott  
TITLE OF INVENTION: INDUCIBLE HERBICIDE RESISTANCE  
NUMBER OF SEQUENCES: 111  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: No. 5942662artis Corporation  
STREET: 3054 Cornwallis Road, P.O. Box 12257  
CITY: Research Triangle Park  
STATE: NC  
COUNTRY: USA  
ZIP: 27709  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/971,217  
FILING DATE:  
CLASSIFICATION: 800  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/457,364  
FILING DATE: 31-MAY-1995  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/181,271  
FILING DATE: 13-JAN-1994  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/093,301  
FILING DATE: 16-JUL-1993  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/937,197  
FILING DATE: 6-NOV-1992  
PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 07/678,378  
FILING DATE: 1-APR-1991  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/305,566  
FILING DATE: 6-FEB-1989  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/165,667  
FILING DATE: 8-MAR-1988  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/042,847  
FILING DATE: 6-APR-1993  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/632,441  
FILING DATE: 21-DEC-1990  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/425,504  
FILING DATE: 20-OCT-1989  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/848,506  
FILING DATE: 6-MAR-1992  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/768,122  
FILING DATE: 27-SEP-1991  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/580,431  
FILING DATE: 7-SEP-1990  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/368,672  
FILING DATE: 20-JUN-1989  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/329,018  
FILING DATE: 24-MAR-1989  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/045,957  
FILING DATE: 12-APR-1993  
ATTORNEY/AGENT INFORMATION:  
NAME: Meigs, J. Timothy  
REGISTRATION NUMBER: 38,241  
REFERENCE/DOCKET NUMBER: S-19825/PL/CCC 1727/DIV5/CONT  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (919)541-8689  
TELEFAX: (919)541-8587  
INFORMATION FOR SEQ ID NO: 103:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 653 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: DNA (genomic)  
US-08-971-217-103

Alignment Scores:  
Pred. No.: 1.29e-45 Length: 653  
Score: 392.00 Matches: 72  
Percent Similarity: 83.04% Conservative: 21  
Best Local Similarity: 64.29% Mismatches: 18  
Query Match: 63.43% Indels: 1  
DB: 2 Gaps: 0

US-09-786-715-8 (1-123) x US-08-971-217-103 (1-653)

OY 5 GUGUGUGLGNValIILeGlyValHsThrValAspGluTrpLysLeuGlnLeuGlnasn 24  
|||||  
DB 75 GAGGAGGACACAGTGTTCGCTGCCACACAGTTGGAGATGGACGAGTACTTCAAGAAA 134  
OY 25 AlalysAspSerLysLysLeuIleValAlAspPheThrAlaSerTrpCysGlyProCys 44  
|||||  
DB 135 GCCGTTGAGACTAAAGAACTGGTGGTGCATTTTACTGCTCATGTGGCGGSCCTTGC 194  
OY 45 ArgPheMetAlaProValLeuAlaGluIleAlaLysLysThrProGluLeuIlePheLeu 64  
|||||  
DB 195 GCTTTATTGCCCCCAATTCCTGCTGACATTCCTAAGAAAGATGCCCATGTATTATTCCTC 254

Qy 65 LysValAspValAspGluValArgProValAla-GluGluTyrSerIleGluAlaMetPr 84  
Db 255 AAGGTGATGTTGATGAACTGAGACGTCTTCCAGCGGAAATGAGTGTGGAGCAATGCC 314  
Qy 84 othrPheLeuPheLeuLysAspGlyGluIleValAspLysValAlaGlyLaserLysAs 104  
Db 315 AACCTTTGCTTCATTAAAGATGGAAGAAAGAGTGGACAGATGTTGCTCCAGAAAGA 374  
Qy 104 PASPLEUGLNAIAThrIleAlaLysHisLaser 115  
Db 375 GGAAGTTCACACAGCCATAGTGAAGCATGCTGCT 408  
RESULT 15  
US-09-350-600-103  
Sequence 103, Application US/09350600  
Patent No. 6262342  
GENERAL INFORMATION:  
APPLICANT: Meins, Frederick  
APPLICANT: Shinshi, Hideaki  
APPLICANT: Menzler, Herman  
APPLICANT: Hofsteenge, Jan  
APPLICANT: Ryals, John  
APPLICANT: Sperisen, Christoph  
TITLE OF INVENTION: DNA SEQUENCES ENCODING POLYPEPTIDES  
TITLE OF INVENTION: HAVING BETA-1,3-GLUCANASE ACTIVITY  
NUMBER OF SEQUENCES: 111  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: No. 6262342artis Corporation  
STREET: 3054 Cornwalis Road, P.O. Box 12257  
CITY: Research Triangle Park  
STATE: NC  
COUNTRY: USA  
ZIP: 27709  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/350,600  
FILING DATE:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/971,217  
FILING DATE: 14-NOV-1997  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/457,364  
FILING DATE: 31-MAY-1995  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/181,271  
FILING DATE: 13-JAN-1994  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/093,301  
FILING DATE: 16-JUL-1993  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/937,197  
FILING DATE: 6-NOV-1992  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/678,378  
FILING DATE: 1-APR-1991  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/305,566  
FILING DATE: 6-FEB-1989  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/165,667  
FILING DATE: 8-MAR-1988  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/042,847  
FILING DATE: 6-APR-1993  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/632,441  
FILING DATE: 21-DEC-1990  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/425,504

FILING DATE: 20-OCT 1989  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/848,506  
FILING DATE: 6-MAR-1992  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/768,122  
FILING DATE: 27-SEP-1991  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/580,431  
FILING DATE: 7-SEP-1990  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/368,672  
FILING DATE: 20-JUN-1989  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/329,018  
FILING DATE: 24-MAR-1989  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/381,443  
FILING DATE: 18-JUL-1989  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/353,312  
FILING DATE: 17-MAY-1989  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/226,303  
FILING DATE: 29-JUL-1988  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/045,957  
FILING DATE: 12-APR-1993  
ATTORNEY/AGENT INFORMATION:  
NAME: Meigs, J. Timothy  
REGISTRATION NUMBER: 38,241  
REFERENCE/DOCKET NUMBER: S-198250  
TELEPHONE: (919)541-8587  
TELEFAX: (919)341-8689  
INFORMATION FOR SEQ ID NO: 103:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 653 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: DNA (genomic)  
US-09-350-600-103  
Alignment Scores:  
Pred. No.: 1,29e-45 length: 653  
Score: 392.00 Matches: 72  
Percent Similarity: 83.04% Conservative: 21  
Best Local Similarity: 64.29% Mismatches: 18  
Query Match: 63.43% Indels: 1  
DB: 3 Gaps: 0  
US-09-786-715-8 (1-123) x US-09-350-600-103 (1-653)  
Qy 5 GluGluGlyGluValIleGlyValHisThrValAspGluTyrPylsLeuGluAlaAsn 24  
Db 75 GAGGAGGAGCAAGTGTGGCGCTCCACAAAGTTGAGGAATGGAACAGTACTTCAAGAAA 124  
Qy 25 AlAlAspSerLysLysLeuIleValAlaAspPheThrLaserTyrPylsGlyProCys 44  
Db 135 GCGGTGAGACTAGAAAGTGGTGGTGGATTTACTGCTTCACTGATGGGSCCTTGC 194  
Qy 45 ArgPheMetAlaProValLeuAlaGluIleAlaLysLysThrProGluLeuIlePheLeu 64  
Db 195 CGTTATATGCCCAATCTCTGCTGACATGCTAAGAAGATGCCCATGTATATTCCTC 254  
Qy 65 LysValAspValAspGluValArgProValAla-GluGluTyrSerIleGluAlaMetPr 84  
Db 255 AAGGTGATGTTGATGAACTGAGACGTCTTCCAGCGGAAATGAGTGTGGAGCAATGCC 314  
Qy 84 othrPheLeuPheLeuLysAspGlyGluIleValAspLysValAlaGlyLaserLysAs 104  
Db 315 AACCTTTGCTTCATTAAAGATGGAAGAAAGAGTGGACAGATGTTGCTCCAGAAAGA 374

QY 104 paspleuGlnAlaThrIleAlaIysHisAlaSer 115  
:::||||| | ||| | |||||:::  
Db 375 GGAGTTCACACACCATAGTGAAGCATGCTGCT 408

Search completed: August 17, 2003, 21:45:28  
Job time : 51.5839 secs



GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

## OM protein - nucleic search, using frame\_plus\_p2n model

Run on: August 17, 2003, 19:25:47 ; Search time 143.373 Seconds  
(without alignments)  
1919.938 Million cell updates/sec

Title: US-09-786-715-8

Perfect score: 618

Sequence: 1 MAEEVEGCVIGVHTDEWKL.....DLQATIAKHAASVAASSS 123

Scoring table: BLOSUM62  
Xgapop 10.0 , Xgapext 0.5  
Ygapop 10.0 , Ygapext 0.5  
Fgapop 6.0 , Fgapext 7.0  
Delop 6.0 , Delext 7.0

Searched: 1504479 seqs, 1118970152 residues

Total number of hits satisfying chosen parameters: 3008958

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 08  
Maximum Match 1008  
Listing first 45 summaries

Command line parameters:  
-MODEL=frame+ p2n model -DEV=xlp  
-Q=/cgn2\_1/USPTO/US09786715/runat\_11082003\_150515\_6129/app\_query.fasta\_1.1052  
-DB=Published\_Applications\_NA -QPM=fastlap -SUFFIX=rmpb -MINMATCH=0.1  
-LOOPEXT=0 -LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blosum62  
-TRANS=human40.cdi -LIST=45 -DOCLALIGN=200 -THR SCORE=pct -THR MAX=100  
-THR MIN=0 -ALIGN=15 -MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEADSIZE=500 -MINLEN=0  
-MAXLEN=2000000000 -USRR=US09786715 @CGN\_1.1.560 @runat\_11082003\_150515\_6129  
-NCPU=6 -ICPU=3 -NO\_MMAP -LARGEIOBERT -NEG\_SCORES=0 -WAIT -DSBLOCK=100  
-LONGLOG -DEV\_TIMEOUT=120 -WARN\_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5  
-FCADPP=6 -FCAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : Published Applications\_NA:  
1: /cgn2\_6/ptodata/2/pubpna/US07\_PUBCOMB.seq:\*  
2: /cgn2\_6/ptodata/2/pubpna/PTC\_NEW\_PUB.seq:\*  
3: /cgn2\_6/ptodata/2/pubpna/US06\_NEW\_PUB.seq:\*  
4: /cgn2\_6/ptodata/2/pubpna/US06\_PUBCOMB.seq:\*  
5: /cgn2\_6/ptodata/2/pubpna/US07\_NEW\_PUB.seq:\*  
6: /cgn2\_6/ptodata/2/pubpna/PTCUS\_PUBCOMB.seq:\*  
7: /cgn2\_6/ptodata/2/pubpna/US08\_NEW\_PUB.seq:\*  
8: /cgn2\_6/ptodata/2/pubpna/US08\_PUBCOMB.seq:\*  
9: /cgn2\_6/ptodata/2/pubpna/US09A\_PUBCOMB.seq:\*  
10: /cgn2\_6/ptodata/2/pubpna/US09B\_PUBCOMB.seq:\*  
11: /cgn2\_6/ptodata/2/pubpna/US09C\_PUBCOMB.seq:\*  
12: /cgn2\_6/ptodata/2/pubpna/US09\_NEW\_PUB.seq:\*  
13: /cgn2\_6/ptodata/2/pubpna/US10A\_PUBCOMB.seq:\*  
14: /cgn2\_6/ptodata/2/pubpna/US10B\_PUBCOMB.seq:\*  
15: /cgn2\_6/ptodata/2/pubpna/US10\_NEW\_PUB.seq:\*  
16: /cgn2\_6/ptodata/2/pubpna/US60\_NEW\_PUB.seq:\*  
17: /cgn2\_6/ptodata/2/pubpna/US60\_PUBCOMB.seq:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	565	91.4	390	US-09-878-574-1462	Sequence 1462, Ap

2	515	83.3	392	10	US-09-878-574-104	Sequence 104, App
3	425	68.8	402	10	US-09-878-574-3723	Sequence 3723, Ap
4	413	66.8	270	10	US-09-878-574-13445	Sequence 13445, A
5	411	66.5	3888	9	US-09-897-898-10	Sequence 10, Appl
6	410	66.3	3425	10	US-09-938-842A-163	Sequence 163, App
7	410	66.3	3149	9	US-09-897-898-5	Sequence 5, Appl1
8	410	66.3	3888	9	US-09-938-842A-164	Sequence 7, Appl1
9	406	65.7	540	12	US-10-349-782-12	Sequence 5, Appl1
10	392	63.4	393	12	US-10-091-841-5	Sequence 12, Appl
11	389	62.9	369	12	US-10-194-885-8	Sequence 5, Appl1
12	389	62.9	369	12	US-10-091-841-1	Sequence 8, Appl1
13	388	62.8	382	12	US-10-091-841-3	Sequence 1, Appl1
14	380.5	61.6	360	10	US-09-938-842A-164	Sequence 164, App
15	377	61.0	659	12	US-10-349-782-9	Sequence 9, Appl1
16	374	60.5	357	10	US-09-938-842A-1736	Sequence 1736, Ap
17	374	60.5	357	14	US-10-322-362-1	Sequence 1, Appl1
18	364	58.9	260	10	US-09-878-574-12941	Sequence 12941, A
19	361	58.4	267	10	US-09-878-574-13137	Sequence 13137, A
20	340	55.0	272	10	US-09-878-574-12181	Sequence 12181, A
21	302	48.9	297	12	US-10-349-782-11	Sequence 11, Appl1
22	298	48.2	277	10	US-09-878-574-8135	Sequence 8135, Ap
23	286	46.3	370	10	US-09-878-574-3169	Sequence 3169, Ap
24	279	45.1	250	9	US-09-923-876-1288	Sequence 1288, Ap
25	275	44.5	268	10	US-09-878-574-5995	Sequence 5995, Ap
26	268	43.4	402	10	US-09-938-842A-225	Sequence 225, App
27	268	43.4	613	9	US-09-770-149-855	Sequence 855, App
28	255	41.3	456	10	US-09-924-035A-677	Sequence 677, App
29	254	41.1	390	10	US-09-938-842A-2688	Sequence 2688, Ap
30	228	36.9	440	9	US-09-770-444-914	Sequence 914, App
31	225	36.4	357	10	US-09-938-842A-147	Sequence 147, App
32	224	36.2	413	14	US-10-102-524-1396	Sequence 1396, Ap
33	224	36.2	418	10	US-09-796-692-3641	Sequence 3641, Ap
34	224	36.2	418	14	US-10-040-862-3641	Sequence 3641, Ap
35	224	36.2	464	10	US-09-998-558-1059	Sequence 1059, Ap
36	224	36.2	479	11	US-09-535-459-1145	Sequence 1145, Ap
37	224	36.2	536	14	US-10-066-543-65	Sequence 65, Appl
38	224	36.2	540	10	US-09-920-300A-1247	Sequence 1247, Ap
39	224	36.2	540	13	US-10-033-528-1247	Sequence 1247, Ap
40	224	36.2	557	10	US-09-884-441-88	Sequence 88, Appl
41	224	36.2	557	11	US-09-907-969-88	Sequence 88, Appl
42	224	36.2	557	14	US-10-198-053-88	Sequence 88, Appl
43	224	36.2	568	11	US-09-535-459-1139	Sequence 1139, Ap
44	224	36.2	570	11	US-09-535-459-1149	Sequence 1149, Ap
45	224	36.2	594	10	US-09-884-441-87	Sequence 87, Appl

## ALIGNMENTS

## RESULT 1

US-09-878-574-1462  
Sequence 1462, Application US/09878574  
Patent No. US20020110548A1

## GENERAL INFORMATION:

APPLICANT: Byrum, Joseph R.  
APPLICANT: La Rosa, Thomas J.  
APPLICANT: Thompson, Michael D.

TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated with

FILE REFERENCE: 38-21(15401)B  
CURRENT APPLICATION NUMBER: US/09/878,574

PRIOR FILING DATE: 2001-12-21  
PRIOR APPLICATION NUMBER: 09/333,535

NUMBER OF SEQ ID NOS: 15775

SEQ ID NO 1462

LENGTH: 390

TYPE: DNA

ORGANISM: Glycine max  
OTHER INFORMATION: Clone ID: LIB3028-037-Q1-B1-F11

US-09-878-574-1462  
Alignment Scores: 1.51e-74 Length: 390



```
RESULT 4
US-09-878-574-13445
; Sequence 13445, Application US/09878574
; Patent No. US20020110548A1
; GENERAL INFORMATION:
; APPLICANT: Byrum, Joseph R.
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Thompson, Michael D.
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated with
; TITLE OF INVENTION: Plants
; FILE REFERENCE: 38-21(15401)B
; CURRENT APPLICATION NUMBER: US/09/878,574
; CURRENT FILING DATE: 2001-12-21
; PRIOR APPLICATION NUMBER: 09/333,535
; PRIOR FILING DATE: 1999-06-14
; NUMBER OF SEQ ID NOS: 15775
; SEQ ID NO 13445
; LENGTH: 270
; TYPE: DNA
; ORGANISM: Glycine max
; OTHER INFORMATION: Clone ID: 701067118H1
US-09-878-574-13445

Alignment Scores:
Pred. No.: 3,42e-52 Length: 270
Score: 413.00 Matches: 80
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 66.83% Indels: 0
DB: 10 Gaps: 0

US-09-786-715-8 (1-123) x US-09-878-574-13445 (1-270)
QY 1 MetAlaGluValGluGluGluValIleGlyValHisThrValAspGluTrpLysLeu 20
DB 29 ATGGCTGAAGTGGAAAGGACAGCGTCATCGGGTCACACCGTGTATGAGGAGCTG 88
QY 21 GlnLeuGlnAsnAlaLysAspSerLysLeuIleValValAspPheThrIleAspTrp 40
DB 89 CAAGCTCCAAATGCAAAACACTCCAAAAACTGATGTGGGATTTACTGCTTCTG 148
QY 41 CysGlyProCysArgPheMetAlaProValIleuAlaGluIleAlaLysThrProGlu 60
DB 149 TGTGTCATGCGCGTTTATGGCCAGTCTTGCGAGATTCGCAAAAGAACTCTGAA 208
QY 61 LeuIlePheLeuLysAlaAspValAspGluValArgProValAlaGluGluTrpSerIle 80
DB 209 TTGATCTTCTCAAAAGTGATGTGAGTGAAGGCGCTGTGCTGAGAAATATTCATT 268

RESULT 5
US-09-897-898-10
; Sequence 10, Application US/09897898
; Patent No. US20020037303A1
; GENERAL INFORMATION:
; APPLICANT: DECKERS, HARM M.
; APPLICANT: VAN ROOIJEN, GIJS
; APPLICANT: BOOTHE, JOSEPH
; APPLICANT: GOUL, JANIS
; APPLICANT: MOLONEY, MAURICE M.
; APPLICANT: DALMIA, BIPIN K.
; TITLE OF INVENTION: THIOREDOXIN AND THIOREDOXIN REDUCTASE CONTAINING OIL
; FILE REFERENCE: 034547/0104
; CURRENT APPLICATION NUMBER: US/09/897,898
; CURRENT FILING DATE: 2001-09-21
; PRIOR APPLICATION NUMBER: 09/577,147
; PRIOR FILING DATE: 2000-05-24
; PRIOR APPLICATION NUMBER: 09/448,600
; PRIOR FILING DATE: 1999-11-24
; PRIOR APPLICATION NUMBER: 09/084,777
; PRIOR FILING DATE: 1998-05-27
; PRIOR APPLICATION NUMBER: 60/047,753
; PRIOR FILING DATE: 1997-05-27
```

```
; PRIOR APPLICATION NUMBER: 60/047,779
; PRIOR FILING DATE: 1997-05-28
; PRIOR APPLICATION NUMBER: 60/075,863
; PRIOR FILING DATE: 1998-02-25
; PRIOR APPLICATION NUMBER: 60/075,864
; PRIOR FILING DATE: 1998-02-25
; NUMBER OF SEQ ID NOS: 24
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 10
; LENGTH: 3888
; TYPE: DNA
; ORGANISM: Unknown Organism
; FEATURE:
; NAME/KEY: CDS
; NAME/KEY: (1555)..(2250)
; LOCATION: (2491)..(2655)
; OTHER INFORMATION: Description of Unknown Organism: Phaseolin
; OTHER INFORMATION: promoter-Trxh oleosin-phaseolin terminator
US-09-897-898-10

Alignment Scores:
Pred. No.: 3,36e-50 Length: 3888
Score: 411.00 Matches: 74
Percent Similarity: 81.90% Conservative: 21
Best Local Similarity: 63.79% Mismatches: 21
Query Match: 66.50% Indels: 0
DB: 9 Gaps: 0

US-09-786-715-8 (1-123) x US-09-897-898-10 (1-3888)
QY 5 GluGluGluGlnValIleGlyValHisThrValAspGluTrpLysLeuGlnAsn 24
DB 1564 GAAGAAGGACAAATGTATCGCGTCACACCGTGAACAATGAGAGCGCTCAGAG 1623
QY 25 AlaLysAspSerLysLysLeuIleValValAspPheThrAlaSerTrpCysGlyProCys 44
DB 1624 GCTAATGATCCAAACTCTTGTGCGTGTGATTACACGCTCTTGCTGTGACCATG 1683
QY 45 ArgPheMetAlaProValLeuAlaGluIleAlaLysThrProGluLeuIlePheLeu 64
DB 1684 CGTTTCATGCGCTCATCTTGTGCTGATTTGGCTAAGAACTCTTAACGCTTTCCTC 1743
QY 65 ThrPheLeuPheLeuLysAspGluGluIleValAspLysValAlaGluLysAsp 84
DB 1744 AAGGTGATCTGATCAATTAAGTCGCGCAAGTATGGCGATACAGCGATGCCA 1803
QY 85 ThrPheLeuPheLeuLysAspGluGluIleValAspLysValAlaGluLysAsp 104
DB 1804 ACCTTCATGTTTGAAGAGAGGAGATTTTGGACAAAGTGTGAGCGCAAGAAAGAT 1863
QY 105 AspLeuGlnAlaThrIleAlaLysHisAlaSerAlaValAlaAla 120
DB 1864 GAGCTTCAGTCTACATTCGCAAAACACTTGGCTATGCGGATACAGT 1911

RESULT 6
US-09-938-842A-163
; Sequence 163, Application US/09938842A
; Patent No. US20020160378A1
; GENERAL INFORMATION:
; APPLICANT: Harper, Jeff
; APPLICANT: Kreps, Joel
; APPLICANT: Wang, Xun
; APPLICANT: Zhu, Tong
; TITLE OF INVENTION: STRESS-REGULATED GENES OF PLANTS, TRANSGENIC PLANTS CONTAINING
; FILE REFERENCE: SCRIPI300-3
; CURRENT APPLICATION NUMBER: US/09/938,842A
; CURRENT FILING DATE: 2001-08-24
; PRIOR APPLICATION NUMBER: US 60/227,866
; PRIOR FILING DATE: 2000-08-24
; PRIOR APPLICATION NUMBER: US 60/264,647
; PRIOR FILING DATE: 2001-01-16
```





PRIOR APPLICATION NUMBER: 60/075,864  
PRIOR FILING DATE: 1998-02-25  
NUMBER OF SEQ ID NOS: 24  
SOFTWARE: Patent In Ver. 2.1  
SEQ ID NO 7  
LENGTH: 3888  
TYPE: DNA  
ORGANISM: Unknown Organism  
FEATURE:  
NAME/KEY: CDS  
LOCATION: (1555)..(1908)  
NAME/KEY: CDS  
LOCATION: (2149)..(2655)  
OTHER INFORMATION: Description of Unknown Organism: Phaseolin  
OTHER INFORMATION: promoter-oleosin Trxn-phaseolin terminator  
US-09-897-898-7

## Alignment Scores:

Pred. No.:	4,73e-50	Length:	3888
Score:	410.00	Matches:	73
Percent Similarity:	85.32%	Conservative:	20
Best Local Similarity:	66.97%	Mismatches:	16
Query Match:	66.34%	Indels:	0
DB:	9	Gaps:	0

US-09-786-715-8 (1-123) x US-09-897-898-7 (1-3888)

QY 5 GUGUGUGYGLINValIIegLYValHISThValAspGLuTrpLysLeuGlnLeuGlnAsn 24  
|||||  
DB 2323 GAAGGAGGACAAAGTGTGAGTGGCTGCGTGCACACCGTTGAGACATGCAAGCAAGCAAG 2382  
25 AAlAspSerLysLysLeuIleValIAspPheThrAlaSerTrpCysGlyProCys 44  
|||||  
QY 2383 GCAATAGAAACCAAACTCTGT 2442  
45 ArgPheMetAlaProValIleuAlaGluIleAlaLysLysThrProGluLeuIlePheLeu 64  
|||||  
DB 2443 CGTTTCATCCCTCATC 2502  
QY 65 LysValAspValAspGluValArgProValAlaGluGluTrpSerIleGluAlaMetPro 84  
|||||  
DB 2503 AAGGTTGATACATGATGAATGAGTCGTCGCAAGTCAATGGCGATACAGCGCATGCCA 2562  
QY 85 ThPheLeuPheLeuLysAspGlyGluIleValIAspLysValAlaGlyAlaSerLysAsp 104  
|||||  
DB 2563 ACCTTCATGTTTGTGAAGGAGGAGATTTTGACAAAGTTGTGGAGCCAAAGAAAGAT 2622  
QY 105 AspLeuGlnAlaThrIleAlaLysHis 113  
|||||  
DB 2623 GAGCTTCAGTCTACCATTTGCCAAACAC 2649

## RESULT 9

US-10-349-782-12  
Sequence 12, Application US/10349782  
Publication No. US20030143618A1  
GENERAL INFORMATION:  
APPLICANT: Yves Hatzfield  
APPLICANT: Valerie Marie-No. US20030143618A111e Frankard  
APPLICANT: Anne-Marie Droual  
TITLE OF INVENTION: Method for easy cloning and selection of chimeric DNA molecules  
FILE REFERENCE: 1187-15  
CURRENT APPLICATION NUMBER: US/10/349,782  
PRIOR FILING DATE: 2003-01-23  
PRIOR APPLICATION NUMBER: EP 02075373.7  
NUMBER OF SEQ ID NOS: 12  
SOFTWARE: Patent In version 3.1  
SEQ ID NO 12  
LENGTH: 540  
TYPE: DNA  
ORGANISM: Artificial Sequence  
FEATURE:  
OTHER INFORMATION: Modified thioredoxin of Oryza sativa in vector pDONR201

US-10-349-782-12

Alignment Scores:  
Pred. No.: 1.03e-50  
Score: 406.00  
Percent Similarity: 78.99%  
Best Local Similarity: 63.03%  
Query Match: 65.70%  
DB: 12  
Gaps: 0

US-09-786-715-8 (1-123) x US-10-349-782-12 (1-540)

QY 5 GUGUGUGYGLINValIIegLYValHISThValAspGLuTrpLysLeuGlnLeuGlnAsn 24  
|||||  
DB 67 GAGAGGAGACTCGTGAATGCTGCTGCGCAACAGAGAGATTCGACGCCAGATACCAAG 126  
25 AAlAspSerLysLysLeuIleValIAspPheThrAlaSerTrpCysGlyProCys 44  
|||||  
DB 127 GCCAGAGAGGCGGCAAAAGTGTATATGACTTCACCTCTCTCTCTCTCTCTCTCTC 186  
45 ArgPheMetAlaProValIleuAlaGluIleAlaLysLysThrProGluLeuIlePheLeu 64  
|||||  
DB 187 CGCTTCATCCCGCCAGTGTGCTGGAATACGCCAAAGTTCCCTGCTGCTCTCTCTC 246  
QY 65 LysValAspValAspGluValArgProValAlaGluGluTrpSerIleGluAlaMetPro 84  
|||||  
DB 247 AAGGTTGATGTTGATGAGTGAAGAGAGAGTGTGAAGATGTCGAGCGCAATGCC 306  
QY 85 ThPheLeuPheLeuLysAspGlyGluIleValIAspLysValAlaGlyAlaSerLysAsp 104  
|||||  
DB 307 ACCTTCATTCATCAAGATGT 366  
QY 105 AspLeuGlnAlaThrIleAlaLysHisAlaSerLysValAlaAlaAlaSerSer 123  
|||||  
DB 367 GACCTCCAGAACACCATGTCGAGGAGCGTGTGCGCCTGCTGATCTGCTTCTGCC 423

## RESULT 10

US-10-091-841-5  
Sequence 5, Application US/10091841  
Publication No. US20030150010A1  
GENERAL INFORMATION:  
APPLICANT: Cho, Myeong-Je  
APPLICANT: Del Val, Greg  
APPLICANT: Calliau, Maxime  
APPLICANT: Lemaux, Peggy G.  
APPLICANT: Buchanan, Bob B.  
TITLE OF INVENTION: Barley Gene for Thioredoxin and  
FILE REFERENCE: 2001-0701.30  
CURRENT APPLICATION NUMBER: US/10/091,841  
PRIOR FILING DATE: 2002-03-05  
PRIOR APPLICATION NUMBER: 09/540,014  
PRIOR FILING DATE: 2000-03-31  
PRIOR APPLICATION NUMBER: US 60/127,198  
PRIOR FILING DATE: 1999-03-31  
PRIOR APPLICATION NUMBER: US 60/169,162  
PRIOR FILING DATE: 1999-12-06  
PRIOR APPLICATION NUMBER: US 60/177,740  
PRIOR FILING DATE: 2000-01-21  
PRIOR APPLICATION NUMBER: US 60/177,739  
PRIOR FILING DATE: 2000-01-21  
NUMBER OF SEQ ID NOS: 51  
SOFTWARE: FastSeq for Windows Version 4.0  
SEQ ID NO 5  
LENGTH: 393  
TYPE: DNA  
ORGANISM: Artificial Sequence  
FEATURE:  
OTHER INFORMATION: wheat thioredoxin h cDNA  
US-10-091-841-5

## Alignment Scores:

Pred. No.:	7.82e-49	Length:	393
------------	----------	---------	-----

Score:	392.00	Matches:	65
Percent Similarity:	82.61%	Conservative:	30
Best Local Similarity:	56.52%	Mismatches:	20
Query Match:	63.43%	Indels:	0
DB:	12	Gaps:	0

US-09-786-715-8 (1-123) x US-10-091-841-5 (1-393)

[illegible]

RESULT 11  
MS-10-194-885-8

```

Sequence 85, Application US/10194885
Publication No. US2003015878A1
GENERAL INFORMATION:
APPLICANT: Wong, J. H.
APPLICANT: Cho, Myeong-je
APPLICANT: Lemaux, Peggy G.
APPLICANT: Buchanan, Bob
TITLE OF INVENTION: TRANSGENIC PLANTS WITH ELEVATED
TITLE OF INVENTION: THIOREDOXIN LEVELS
FILE REFERENCE: 416212000800
CURRENT APPLICATION NUMBER: US/10/194,885
CURRENT FILING DATE: 2002-07-12
PRIOR APPLICATION NUMBER: 60/307,006
PRIOR FILING DATE: 2001-07-19
PRIOR APPLICATION NUMBER: 09/538,864
PRIOR FILING DATE: 2000-03-29
PRIOR APPLICATION NUMBER: 60/126,736
PRIOR FILING DATE: 1999-03-29
NUMBER OF SEQ ID NOS: 55
SOFTWARE: FastSeq for Windows Version 4.0

```

```

? SEQ ID NO 8
? LENGTH: 369
? TYPE: DNA
? ORGANISM: Barley
? FEATURE:
? NAME/KEY: CDS
? LOCATION: (1)...(365)
US-10-194-885-8

Alignment Scores:
Pred. No.:      1 99e-48
Score:          389.00
Percent Similarity: 85.32%
Best Local Similarity: 58.72%
Query Match:    62.94%
DB:             12

Length: 365
Matches: 64
Conservative: 29
Mismatches: 16
Indels: 0
Gaps: 0

```

US-09-786-715-8 (1-123) x US-10-194-885-8 (1-369)

QY 8 GlnValIleGlyValHisThrValAspGluTrpLysLeuGlnLeuGlnAsnAlaLysAsp 27  
Db 37 GAGGTGATCTCGSTCCACAGCCTGGAGCACTGGACCATTGCAGATCGAGAGGCCAACACC 96  
QY 28 SerLysLysLeuIleValValAlaSpherThrAlaSerTrpCysGlyProCysArgPheMet 47  
Db 97 GCCAAAGACGTGGTGATGTGATTACCTTCACGTGATCATGTGTGGGACCATGGCGCATG 156  
QY 48 AlaProValLeuAlaGluIleAlaLysLysThrProGluLeuIlePheLeuLysValAsp 67  
Db 157 GCTCCAGTTTGGCTGATCTCGCCCAAGAGTTCCCAATGCTGTTTCCTCACAGTGAC 216  
QY 68 ValAspGluValArgProValAlaGluGluTrpSerIleGluAlaMetProThrPheLeu 87  
Db 217 GTGGATGAAGACTGAAGCCCATTTGCTAGCAATTCACATGTCGAGGCCATGCAACGCTTCGTG 276  
QY 88 PheLeuLysAspGlyValIleValAlaSplysValValGlyAlaSerLysAspAspLeuGln 107  
Db 277 TTGATTAAGAGAGAGACGTCAAGGACAGGGTTGTGGAGCTATTCAAAGAGAGACTGACC 336  
QY 108 AlaThrIleAlaLysHisAlaSerAla 116  
Db 337 GCCAAGGTTGGCTTCACGCGCGGCC 363

RESULT 12  
US-10-091-841-1  
; Sequence 1, Application US/10091841  
; Publication No. US20030150010A1

```

: GENERAL INFORMATION:
: APPLICANT: Cho, Myeong-Je
: APPLICANT: Del Val, Greg
: APPLICANT: Calliau, Maxime
: APPLICANT: Lemauz, Peggy G.
: APPLICANT: Buchanan, Bob B.
: TITLE OF INVENTION: Barley Gene for Thioredoxin and
: TITLE OF INVENTION: NADP-Thioredoxin Reductase
: FILE REFERENCE: 2001-0701.30
: CURRENT APPLICATION NUMBER: US/10/091,841
: CURRENT FILING DATE: 2002-03-05
: PRIOR APPLICATION NUMBER: 09/540,014
: PRIOR FILING DATE: 2000-03-31
: PRIOR APPLICATION NUMBER: US 60/127,198
: PRIOR FILING DATE: 1999-03-31
: PRIOR APPLICATION NUMBER: US 60/169,162
: PRIOR FILING DATE: 1999-12-06
: PRIOR APPLICATION NUMBER: US 60/177,740
: PRIOR FILING DATE: 2000-01-21
: PRIOR APPLICATION NUMBER: US 60/177,739
: PRIOR FILING DATE: 2000-01-21
: NUMBER OF SEQ ID NOS: 51
: SOFTWARE: FastSeq for Windows Version 4.0
: SEQ ID NO 1
: LENGTH: 369
: TYPE: DNA
: ORGANISM: Artificial Sequence
: FEATURE:
: OTHER INFORMATION: barley thioredoxin h cDNA
: US-10-091-841-1

```

Alignment Scores:	
Pred. No.:	1,99e-48
Score:	389.00
Percent Similarity:	85.32%
Best Local Similarity:	58.72%
Query Match:	62.94%
BB:	12
Length:	366
Matches:	64
Conservative:	29
Mismatches:	16
Indels:	0
Gaps:	0

US-09-786-715-8 (1-123) x US-10-091-841-1 (1-369)

```
QY      8 GlnValIleGlyValHisThrValAspLutRplyLeuGlnLeuGlnAsnAlaLysAsp 27
          ::|||||  |||||:::~::~|||  ::|||:::~  ||
Db     37 GAGTGATCTCCGGTCCACAGCCTGGAGCACAGTGCACCATTGAGAGGCCAACACC 96
```

QY	28	leerlyslsleuileValValasphetrAlasertrpcyglyprocyasargPheMet	47
		.....	111
Db	97	GCCAGAGAGCTGGTGATTTGACTTCACGTGATCGTGGGACCATGCCGATCAG	156
QY	48	AlaProValLeuAlaGluIleAlaLysLysThrProGluLeuIlePheLeuLysValasP	67
		.....	111
Db	157	GCTCCAGTTTTCGCGATCTCGCCCAAGAAGTTCCCAATAGTGTTCCTCAAGGTGCAG	216
QY	68	ValaspgIuValArgProValAlaGluGluTyrSerIleGluAlaMetProThrPheLeu	87
		.....	111
Db	217	GTGGGTGACGTCAACCCCATTTGCTGACCAATTGACGTGTGAGGCCATTGCCAAGTTCGTG	276
QY	88	PheLeuLysAspGlyGluIleValasPlysValAlaGlyLaserLysAspSpleuGln	107
		.....	111
Db	277	TTTCATGAAGGAAGGAGACGTCAAGGACAGCGTTGTTCGAGCATATCAAGAGGACACGTACC	336
QY	108	AlaThrIleAlaLysHisAlaSerAla	116
		.....	111
Db	337	GCCAAAGTGTGGCTTCACGCGCGCGCC	363

```

RESULT 13
US-10-091-841-3
; Sequence 3, Application US/10091841
; Publication No. US20030150010A1
; GENERAL INFORMATION:
; APPLICANT: Cho, Myeong-Je
; APPLICANT: Del Val, Greg
; APPLICANT: Calliau, Maxime
; APPLICANT: Lemaux, Peggy G.
; TITLE OF INVENTION: Barley Gene for Thioredoxin and
; FILE REFERENCE: 2001-0701.30
; CURRENT APPLICATION NUMBER: US/10/091,841
; CURRENT FILING DATE: 2002-03-05
; PRIOR APPLICATION NUMBER: 09/540,014
; PRIOR FILING DATE: 2000-03-31
; PRIOR APPLICATION NUMBER: US 60/127,198
; PRIOR FILING DATE: 1999-03-31
; PRIOR APPLICATION NUMBER: US 60/169,162
; PRIOR FILING DATE: 1999-12-06
; PRIOR APPLICATION NUMBER: US 60/177,740
; PRIOR FILING DATE: 2000-01-21
; PRIOR APPLICATION NUMBER: US 60/177,739
; PRIOR FILING DATE: 2000-01-21
; NUMBER OF SEQ ID NOS: 51
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 3
; LENGTH: 382
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: wheat thioredoxin h cDNA
US-10-091-841-3

Alignment Scores:
Pred. NO.:      2.95e-48          Length:      382
Score:          388.00           Matches:     64
Percent Similarity: 82.46%       Conservative: 30
Best local Similarity: 56.14%    Mismatches:  20
Query Match:    62.78%          Indels:      0
DB:              12             Gaps:        0

US-09-786-715-8 (1-123) x US-10-091-841-3 (1-382)

OY      2 A|aglvAlJglunglunGlnValIlleGlyValHtHStrValAsplutRplyslengIn 21
         ||||| ||||| |||||:||||| ||||||:|||||:||||| :|||
Db      37 GCGGGGTAGGGGGGGGGAGGTGATCTCCACACACCTGGAGCAGTGCATGCG 96

OY      22 LeunglnsAlatAlasAspSerIysIysLeuIleValVAlasPheThAlaseTrpCys 41
         ::::: ||| :|||:|||||:|||||:|||||:|||||:|||||
Db      97 ATCAAGAGAGGCCAACGCCGCCACAGACTGTGGTGATTGAATTTCATCATCATGTC 156

```

QY	42	GlyProCysArgPheMetAlaProValLeuAlaGluIleAlaLysLysThrProGluLeu	61
Db	157	GGACCACTGGCGGCAATTAATGCTCCAAATTTTCGTATCTCCGCAAGAAGTTCCACGTCT	216
QY	62	IlePheLeuLysValAlaSpValAspGluValAcgProValAlaGluGluTyrSerLeu	81
Db	217	GTTTCTCTCAAGTCGACGTTTATGTAAGTCAAGCCCATTCGTCGACATTCAGCGTGAG	276
QY	82	AlaMetProThrPheLeuPheLeuLysAspGlyIuIleValAspLysValValGlyAla	101
Db	277	GCCATGCCAACCTTCCTGTTTCATGCAAGAGACAGATGTCAGAGACAGGTTGCGAGCT	336
QY	102	SerLysAspAspLeuGlnAlaThrIleAlaLysHisAlaSer	115
Db	337	ATCAAGAGAGAACTGACGACCAAGATTGGGCTACACGGCGGC	378

```

RESULT 14
US-09-938-842A-164
Sequence 164, Application US/09938842A
Patent No. US20020160378A1
GENERAL INFORMATION:
APPLICANT: Harper, Jeff
APPLICANT: Krepes, Joel
APPLICANT: Wang, Xun
APPLICANT: Zhu, Tong
TITLE OF INVENTION: STRESS-REGULATED GENES OF PLANTS, TRANSGENIC PLANTS CONTAINING SUCH GENES
FILE REFERENCE: SAME, AND METHODS OF USE
CURRENT APPLICATION NUMBER: US/09/938,842A
PRIORITY FILING DATE: 2001-08-24
PRIOR APPLICATION NUMBER: US 60/227,866
PRIOR FILING DATE: 2000-08-24
PRIOR APPLICATION NUMBER: US 60/264,647
PRIOR FILING DATE: 2001-01-16
PRIOR APPLICATION NUMBER: US 60/300,111
PRIOR FILING DATE: 2001-06-22
NUMBER OF SEQ ID NOS: 5379
SEQ ID NO 164
LENGTH: 360
TYPE: DNA
ORGANISM: Arabidopsis thaliana
US-09-938-842A-164

Alignment Scores:
Pred. No.: 3,52e-47 Length: 360
Score: 380.50 Matches: 71
Percent Similarity: 76.72% Conserved: 18
Best Local Similarity: 61.21% Mismatches: 26
Query Match: 61.57% Indels: 1
DB: 10 Gaps: 1

US-09-786-715-8 (1-123) x US-09-938-842A-164 (1-360)

QY      5  GLUGLUGLYGLNVALILEGLIYVAHISThrValAspGluTrpLysLeuGlnLeuGlnAsn 24
       |||||||
Db      10  GAAGAGGGCAACGATGTGTTGCACACGAAACGATGATGAGCTGTGCCAACCTTGATAAA 69
       |||||||

QY     25  ALALYSASPSerLysLysLeuLeuValAlaAspSerThrAlaSerTPySGLyrProCys 44
       |||||||
Db      70  GCCAAAGAATTCACAACACACTGATTGTGATTGAATTAATTCACCTCGTTCATGTGTTCCACCATTGC 129
       |||||||

QY     45  AcArgPheMetAlaProValLeuAlaGluIleAlaLysLys--ThrpGluLeuIlePhe 63
       |||||
Db     130  CCGAGTAGTTCGCCCAATTTTCCACGATTTGGCCAAAGAAAGUATCATGTCMAATGCCATCTTC 189
       |||||

QY     64  LeuLysValAspValAspGluValAlarProValAlaGluLutrySerIleGluAlaMet 83
       |||||||
Db     190  TTCACAGGTGATGTGATGATGACCTTCACAGTGTGCTTAAAGAGTTTGTTGGTGTGAGGCAATG 249
       |||||||

QY     84  ProThrPheLeuPheLeuLysAspGlyGluIleValAspLysValValGlyAlaSerLys 103
       |||||||
Db     250  CCAACCTTGTGTTTATTAAAGCCGGCAACAGTGTGATTAAGCTTCGTTGGTGTGGAATAAA 309

```

Qy 104 ASPASPLEUcInAlaThrIleAlaLysHSAIaSerAlaValAlaAla 119  
:::||||| ||| ||||| |||  
Db 310 GAAGATCTTCAGCGGAAATAAGTGAAGCACTACTGCTTTACAACTCGC 357

### RESULT 15

US-10-349-782-9  
Sequence 9 A

; Sequence 9, Application US/10349782  
; Publication No. US20030143618A1

Publication No. US20030143618A1

GENERAL INFORMATION:

APPLICANT: Yves Hatzfield

APPLICANT: Valerie Marie-No. US20030143618A111e Frankard

APPLICANT: Anne-Marie Droual

7 TITLE OF INVENTION: Method for easy cloning and selection of chimeric DNA molecules

FILE REFERENCE: 1187-15

FILE REFERENCE: 1187-13  
CURRENT APPLICATION NUMBER: D

;; CURRENT APPLICATION NUMBER: US/10/349, /  
: CURRENT FILING DATE: 2003-01-23

CURRENT FILING DATE: 2003-01-21  
PRIOR APPLICATION NUMBER: 00/000000

PRIOR APPLICATION NUMBER

PRIOR FILING DATE: 2002-01-23

NUMBER OF S

; SOFTWARE: Pa

; SEQ.ID NO 9

```

: SEQ ID NO 9
:
: LENGTH: 659

```

LENGTH: 659  
TYPE: DNA

```

;
TYPE: DNA
OCCASION: 2014:01:01 00:00:00

```

ORGANISM: A

### Alignment Scores:

Pred. No.:	2.82e-46	Length:	659
------------	----------	---------	-----

Score:	377.00	Matches:	74
--------	--------	----------	----

Score:	377.00	Matches:	74
Percent Similarity:	78.15%	Conservative:	19

Percent similarity:	78.13%	Conservative:	19
Best local similarity:	62.18%	Mismatches:	26

Best Local Similarity:	62.18%	Mismatches:	26
Query Match:	61.00%	Indels:	1

query Match:	61.00%	Indels:	1
DB:	12	Calls:	0

US-09-786-715-8 (1-123) x US-10-349-782-9 (1-659)

5 G L I U G L I G L I N V A I I E G L Y A H I S H R V A L A S P G L T R P L Y S L E U G L I N L E U G L I N A S N 20

139 GAGGAGGAGGATCGTGTATCGCCCTGCCACAAACAAAGACGAGTTGGACGCCACAGATGACCAAG 199

Db 199 GCCAAGGAGCCCGCAAGTGGCTAATTTAGCTTCACTGCTTCCTGGTGTGGCCCTTGC 255

45 ArgPheMetAlaProValLeuAlaGluIleAlaLysThrProGluLeuIlePheLeu 64

D6 259 CCGTTCATCCGCCACAGTGTTCCGCTGATACGCCAAAAGTCCCTCCCTGGTGTCTCTCTCTG 318

319 AAGGTTGATGTTGATGAGCTGAGCAGAGATTGCTGAAAAGTACATGTCGAGGCAATGCCG 378

Search completed: August 17, 2003, 23:43:55  
Job time : 146.373 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - nucleic search, using frame\_plus\_p2n model

Run on: August 17, 2003, 19:20:27 ; Search time 1663.43 Seconds

(without alignments)  
1797.162 Million cell updates/sec

29: gb\_gss2:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Title: US-09-786-715-8  
Perfect score: 618  
Sequence: 1 MAVEERGQYGVHTVDEMKL.....DDLQATIAKASAVAAASS 123

Scoring table: BLOSUM62  
Xgapop 10.0, Xgapext 0.5  
Ygapop 10.0, Ygapext 0.5  
Fgapop 6.0, Fgapext 7.0  
Delop 6.0, Delext 7.0

Searched: 22781392 seqs, 12152238056 residues

Total number of hits satisfying chosen parameters: 45562784

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Command line parameters:  
-MODEL-frame+pn.model -DEV-xlp  
-Q/cg2\_1/USPTO.spool/US09786715/runat.11082003.150514.6048/app-query.fasta.1.1052  
-DB-EST -OPMT-fastap -SUFFIX=1st -MINMATCH=0.1 -LOOPEXT=0 -LOOPEXT=0  
-UNITS-bits -START=1 -END=1 -MATRIX-blosum62 -TRANS-human40.cdi -LIST=45  
-DOALIGN=200 -THR.SCORE=pct -THR.MAX=100 -THR.MIN=0 -ALIGN=15 -MODE=LOCAL  
-OUTFMT=ptc -NORM=ext -HEADSIZE=500 -MINLEN=0 -MAXLEN=200000000  
-USER=US09786715.ecn.1.1.783-etrnat.11082003.150514.6048 -NCPU=6 -ICPU=3  
-NO\_MMAP -LARGEQUERY -NEG.SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG  
-DEV.TIMEOUT=120 -WARN.TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6  
-FAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :  
EST: \*  
1: em\_estba: \*  
2: em\_esthum: \*  
3: em\_estlin: \*  
4: em\_estmu: \*  
5: em\_estrov: \*  
6: em\_estrpl: \*  
7: em\_estro: \*  
8: em\_hic: \*  
9: gb\_est1: \*  
10: gb\_est2: \*  
11: gb\_hic: \*  
12: gb\_est3: \*  
13: gb\_est4: \*  
14: gb\_est5: \*  
15: em\_estfun: \*  
16: em\_estom: \*  
17: em\_gss\_hum: \*  
18: em\_gss\_inv: \*  
19: em\_gss\_pin: \*  
20: em\_gss\_vrt: \*  
21: em\_gss\_fun: \*  
22: em\_gss\_mam: \*  
23: em\_gss\_mus: \*  
24: em\_gss\_pro: \*  
25: em\_gss\_rod: \*  
26: em\_gss\_png: \*  
27: em\_gss\_vr1: \*  
28: gb\_gss1: \*

Result No.	Score	Query Match	Length	DB ID	Description
1	618	100.0	441	9	AW705063
2	618	100.0	452	10	BE807885
3	618	100.0	455	10	BE473545
4	618	100.0	464	10	BE238189
5	618	100.0	465	9	AW781479
6	618	100.0	475	10	BE347100
7	618	100.0	479	10	BE807019
8	618	100.0	505	10	BE725950
9	618	100.0	508	10	BE607957
10	618	100.0	510	9	AW568753
11	618	100.0	517	9	AW164347
12	618	100.0	527	12	BM884981
13	618	100.0	532	14	CD414498
14	618	100.0	544	10	BE440397
15	618	100.0	559	12	BM177477
16	618	100.0	577	14	CD414001
17	618	100.0	578	14	CD403924
18	618	100.0	584	14	CD406706
19	618	100.0	608	14	CA935518
20	618	100.0	662	13	BU548381
21	614	99.2	519	14	CA937099
22	613	99.2	580	14	CD416637
23	605	97.9	487	9	A1736736
24	600	97.1	497	12	B1941622
25	595	96.3	425	13	BU763193
26	593	96.0	488	9	AW277335
27	576	93.2	372	9	AW597424
28	572	92.6	511	14	CD407464
29	565	91.4	428	10	BC046833
30	538	87.1	349	10	BF068694
31	535	86.6	426	14	CD416748
32	516	83.5	479	12	BC838458
33	516	83.5	534	14	CA799061
34	516	83.5	616	14	CD410224
35	515	83.3	390	9	A1441505
36	515	83.3	547	9	AW349509
37	515	83.3	556	12	B1471125
38	515	83.3	570	14	CD398312
39	515	83.3	612	14	CD487665
40	511	82.7	373	10	BE210731
41	508	82.2	433	13	BO786368
42	471	76.2	350	9	A1461254
43	452	73.1	464	9	AL386021
44	452	73.1	471	9	AW329521
45	452	73.1	478	9	AL384324

## ALIGNMENTS

RESULT 1  
LOCUS AW705063  
DEFINITION  
sk57b09.y1 Gm-c1019 glycine max cDNA clone GENOME SYSTEMS CLONE ID: Gm-c1019-6642 5' similar to SW:THIL\_TOBAC Q07090 THIOREDOXIN H-TYPE 2 ; mRNA sequence.  
ACCESSION AW705063  
VERSION AW705063.1 GI:7569286  
KEYWORDS  
SOURCE ORGANISM  
glycine max (soybean)  
EST.  
glycine max  
glycine max  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids

; eurousids I; Fabales; Fabaceae; Papilionoideae; Phaseoleae; Glycine.  
 1 (bases 1 to 441)  
 Shoemaker, R., Keim, P., Vodkin, L., Erpelting, J., Coryell, V., Khanna, A., Bolla, B., Maria, M., Hillier, L., Kucaba, T., Martin, J., Beck, C., Wylie, T., Underwood, K., Steptoe, M., Theising, B., Allen, M., Bowers, Y., Pearson, B., Swaller, T., Gibbons, M., Pape, D., Harvey, N., Schurk, R., Ritter, E., Kohn, S., Shin, T., Jackson, Y., Cardenas, M., McCann, R., Waterston, R. and Wilson, R.  
 Public Soybean EST Project  
 Unpublished  
 Contact: Shoemaker R/Public Soybean EST Project  
 Public Soybean EST Project  
 Washington University School of Medicine  
 4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA  
 Tel: 314 286 1800  
 Fax: 314 286 1810  
 Email: est@watson.wustl.edu  
 This clone is available through: Resgen, Invitrogen Corp. 2130 South Memorial Parkway Huntville, AL 35801 For further information call: (800)-533-4363 or contact via email: ccluresgen.com  
 Insert Length: 657 Std Error: 0.00  
 Seq primer: -40RP from Gibco  
 High quality sequence stop: 425.  
 Location/Qualifiers  
 1..441  
 /organism="Glycine max"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:3847"  
 /clone="GENOME SYSTEMS CLONE ID: Gm-cl019-6642"  
 /tissue\_type="Immature seed coats of greenhouse grown plants"  
 /lab\_host="DH10B (Gibco BRL)"  
 /note="Vector: pSPORT1 (Life Technologies); Site 1: Not I; Site 2: Sal I; This cDNA library was constructed from mRNA isolated from immature seed coats (200-300 mgs) of greenhouse grown plants. The library was prepared using the Life technologies pSuperScript cDNA library construction kit. Complementary DNA was synthesized from mRNA using a poly (dT) sequence with a Not I restriction site. Sal I linker adapters were ligated to the blunt-ended cDNA fragments followed by Not I digestion. The cDNA fragments were directionally cloned into the Not I-Sal I restriction site of the pSPORT1 vector. The ligated cDNA fragments were transformed into E.coli Electromax DH10B host cells (Gibco BRL). This library was constructed by Dr. Lilla Vodkin and Dr. Anu Khanna."  
 BASE COUNT 118 a 88 c 112 g 123 t  
 ORIGIN  
 Alignment Scores:  
 Pred. No.: 3,31e-75 Length: 441  
 Score: 618.00 Matches: 123  
 Percent Similarity: 100.00% Conservative: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 100.00% Indels: 0  
 DB: 9 Gaps: 0  
 US-09-786-715-8 (1-123) x AM705063 (1-441)  
 QY 1 MetAAGUVALGIUGLUGLIVALLIeGLYVALHISTHVALASpGLUTRPLYSLeu 20  
 Db 27 ATGCTGAAGTGAAGAGGAGAGTCTATCGCGCCACCGCTGATGAGTGGAGAGCTG 86  
 QY 21 GlnLeuGlnAsnAlaLysAspSerLysLysLeuIleValIAspPheThrAlaSerTrp 40  
 Db 87 CAACTCCAGATGCCAAAGAACTCCAAAGAACTGATGGTGATTTACTGCTCCCTGG 146  
 QY 41 CysGLYProCysArgPheMetAlaProValLeuAlaGluIleAlaLysLysThrProGlu 60  
 Db 147 TGTGTGCATGCCGTTTATGTGCGCCCACTTCTTGAGAGATTGCCAAAGAAATCCTGAA 206

QY 61 LeuIlePheLeuLysValAspValAspGluValArgProValAlaGluGluTrpSerIle 80  
 Db 207 TTGATCTTCCTCCAAAGTGGATGGATGAGTGAAGTGAAGCGCTGCTGAGGAATATTCAT 266  
 QY 81 GluAlaMetProThrPheLeuPheLeuLysAspGlyGluIleValAspLysValValGly 100  
 Db 267 GAGGCCATGCCAACCTCTCTCTTGAAGATGGCGAGATCGTGGACAGAGGTGTGCT 326  
 QY 101 AlaSerLysAspAspLeuGlnAlaThrIleAlaLysSHLSAlaSerAlaValAlaAla 120  
 Db 327 GCTACTAAGATGAGACCTTCACGCCACCATAGCAAGCATGATCTGCTGCTCTCTCT 386  
 QY 121 SerSerSer 123  
 Db 387 TCTTCTCT 395  
 RESULT 2  
 BE807885  
 LOCUS  
 DEFINITION  
 Gm-cl061-345 5' similar to SW:TH11\_TOBAC 007090 THIOREDOXIN H-TYPE 2 ; mRNA sequence.  
 BE807885  
 BE807885.1 GI:10238997  
 EST.  
 VERSION  
 KEYWORDS  
 SOURCE  
 ORGANISM  
 Glycine max (soybean)  
 Glycine max  
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids ; eurousids I; Fabales; Fabaceae; Papilionoideae; Phaseoleae; Glycine.  
 1 (bases 1 to 452)  
 Shoemaker, R., Keim, P., Vodkin, L., Erpelting, J., Coryell, V., Khanna, A., Bolla, B., Maria, M., Hillier, L., Kucaba, T., Martin, J., Beck, C., Wylie, T., Underwood, K., Steptoe, M., Theising, B., Allen, M., Bowers, Y., Pearson, B., Swaller, T., Gibbons, M., Pape, D., Harvey, N., Schurk, R., Ritter, E., Kohn, S., Shin, T., Jackson, Y., Cardenas, M., McCann, R., Waterston, R. and Wilson, R.  
 Public Soybean EST Project  
 Unpublished  
 Contact: Shoemaker R/Public Soybean EST Project  
 Public Soybean EST Project  
 Washington University School of Medicine  
 4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA  
 Tel: 314 286 1800  
 Fax: 314 286 1810  
 Email: est@watson.wustl.edu  
 This clone is available through: Resgen, Invitrogen Corp. 2130 South Memorial Parkway Huntville, AL 35801 For further information call: (800)-533-4363 or contact via email: ccluresgen.com  
 High quality sequence stop: 414.  
 Location/Qualifiers  
 1..452  
 /organism="Glycine max"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:3847"  
 /clone="GENOME SYSTEMS CLONE ID: Gm-cl061-345"  
 /tissue\_type="mature flowers of field grown plants"  
 /lab\_host="DH10B"  
 /note="Vector: pBluescript II SK+, Site 1: EcoRI; Site 2: XhoI; The cDNA library was constructed from mRNA isolated from mature flowers of field grown plants for the cultivar Ralden. Complementary DNA was synthesized from mRNA using a primer consisting of a poly(dT) sequence with a XhoI restriction site. EcoRI adapters were ligated to the blunt-ended cDNA fragments followed by XhoI digestion. The cDNA fragments were directionally cloned into the EcoRI-XhoI restriction site of the pBluescript vector. The ligated cDNA fragments were transformed into DH10B host cells (GibcoBRL). This library was constructed in the laboratory of Dr. Randy Shoemaker."  
 BASE COUNT 123 a 87 c 115 g 127 t

## ORIGIN

## Alignment Scores:

Pred. No.:	3.43e-75	Length:	452
Score:	618.00	Matches:	123
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	100.00%	Indels:	0
DB:	10	Gaps:	0

US-09-786-715-8 (1-123) x BE807885 (1-452)

QY 1 MetAlaGluValGluGluGluGluValIleGlyValHisThrValAspGluTrpLysLeu 20  
|||||  
DB 21 ATGGCTGAAGTGAAGAGGACAGCATCGGCTCCACACCTTGATGATGAGACTG 80  
QY 21 GlnLeuGlnAsnAlaLysAspSerLysLysLeuIleValValAspPheThrAlaSerTrp 40  
|||||  
DB 81 CAACCTCAGATGCAAAAGACTCCAAAACATGATGTGTGATTTTACTGCTCTCTG 140  
QY 41 CysGlyProCysArgPheMetAlaProValLeuAlaGluIleAlaLysLysThrProGlu 60  
|||||  
DB 141 TGTGGTCCAGCCGCTTTATAGCCCGAGTTCTTGACAGATTCAGAAAGAACTCTGAA 200  
QY 61 LeuIlePheLeuLysValAspValAspGluValArgProValAlaGluGluTrpSerIle 80  
|||||  
DB 201 TTGATCTTCTCAAAATGATGATGTGATGAGTGAAGCTGTCTCTGAGAAATATTCATT 260  
QY 81 GluAlaMetProThrPheLeuPheLeuLysAspGlyGluIleValAspLysValAlaGly 100  
|||||  
DB 261 GAGGCCATGCGCAACCTTCTCTTGAAGATGCGAGATCTGAGCAAGGTGTGTGT 320  
QY 101 AlaSerLysAspLysLeuGlnAlaThrIleAlaLysHisAlaSerAlaValAlaAla 120  
DB 321 GCTAGTAAGATGATCCTTCAAGCCACCATAGCCCAAGCATGATCTGCTGTGCTGCT 380  
QY 121 SerSerSer 123  
|||||  
DB 381 TCTTCTTCT 389

RESULT 3 BE473545 455 bp mRNA linear EST 04-DEC-2001  
LOCUS BE473545  
DEFINITION sp52d10.y1 Gm-c1043 glycine max cDNA clone GENOME SYSTEMS CLONE ID:  
Gm-c1043-1868 5' similar to SW:TH11\_TOBAC Q07090 THIOREDOXIN H-TYPE  
2' mRNA sequence.

ACCESSION BE473545  
VERSION BE473545.1 GI:9564036  
KEYWORDS EST.  
SOURCE glycine max (soybean)  
ORGANISM Glycine max

REFERENCE Eukaryota: Viridiplantae: Streptophyta: Embryophyta: Tracheophyta;  
Spermatophyta: Magnoliophyta: eudicotyledons: core eudicots: rosids  
; eustroids I; Fabales; Fabaceae; Papilionoideae; Phaseoleae;  
Glycine.  
1 (bases 1 to 455)  
Shoemaker, R., Keim, P., Vodkin, L., Erpelting, J., Corryell, V., Khana  
A., Bolla, B., Matra, M., Hillier, L., Kucada, T., Martin, J., Beck, C.,  
Wille, T., Underwood, R., Steptoe, M., Theising, B., Allen, M., Bowers  
, Y., Person, B., Swaller, T., Gibbons, M., Pape, D., Harvey, N., Schuck  
, R., Ritter, E., Kohn, S., Shin, T., Jackson, Y., Cardenas, M., McCann  
, R., Waterston, R. and Wilson, R.  
Public Soybean EST Project  
unpublished

TITLE JOURNAL  
COMMENT Contact: Shoemaker R./Public Soybean EST Project  
Public Soybean EST Project  
Washington University School of Medicine  
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA  
Tel: 314 286 1800  
Fax: 314 286 1810

Email: est@wustl.edu  
This clone is available through: ResGen, Invitrogen Corp. 2130  
South Memorial Parkway Huntsville, AL 35801 For further information

call: (800)-533-4363 or contact via email: ccu@resgen.com  
Insert length: 596 Std Error: 0.00  
High quality sequence stop: 444.

## FEATURES

## source

Location/Qualifiers

1..455

/organism="Glycine max"

/mol\_type="mRNA"

/db\_xref="taxon:3847"

/clone="GENOME SYSTEMS CLONE ID: Gm-c1043-1868"

/tissue\_type="Hypocotyl and Plumule, germinating seeds"

/lab\_host="DH10B"

/clone\_11b="Gm-c1043"

/note="Vector: pT73Pac (Pharmacia). Site 1: EcoRI;  
Site 2: NotI. This cDNA library was constructed from mRNA  
isolated from hypocotyl and plumule tissues of seeds  
germinated for three days of the cultivar Williams.  
Complementary DNA was synthesized from mRNA using a primer  
consisting of a poly(dT) sequence with a NotI restriction  
site. EcoRI adapters were ligated to the blunt-ended cDNA  
fragments followed by digestion with EcoRI and NotI. The  
cDNA fragments were directionally cloned into the  
EcoRI-NotI restriction site of the pT73-Pac vector. The  
ligated cDNA fragments were transformed into DH10B host  
cells (Gibco BRL). This library was constructed by Dr.  
Randy Shoemaker."

BASE COUNT 125 a 89 c 115 g 125 t 1 others  
ORIGIN

## Alignment Scores:

Pred. No.:	3.46e-75	Length:	455
Score:	618.00	Matches:	123
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	100.00%	Indels:	0
DB:	10	Gaps:	0

US-09-786-715-8 (1-123) x BE473545 (1-455)

QY 1 MetAlaGluValGluGluGluGluValIleGlyValHisThrValAspGluTrpLysLeu 20  
|||||  
DB 34 ATGGCTGAAGTGAAGAGGACAGCATCGGCTCCACACCTTGATGATGAGACTG 93  
QY 21 GlnLeuGlnAsnAlaLysAspSerLysLysLeuIleValValAspPheThrAlaSerTrp 40  
|||||  
DB 94 CAACCTCAGATGCAAAAGACTCCAAAACATGATGTGTGATTTTACTGCTCTCTG 153  
QY 41 CysGlyProCysArgPheMetAlaProValLeuAlaGluIleAlaLysLysThrProGlu 60  
|||||  
DB 154 TGTGGTCCAGCCGCTTTATAGCCCGAGTTCTTGACAGATTCAGAAAGAACTCTGAA 213  
QY 61 LeuIlePheLeuLysValAspValAspGluValArgProValAlaGluGluTrpSerIle 80  
|||||  
DB 214 TTGATCTTCTCAAAATGATGATGTGATGAGTGAAGCTGTCTCTGAGAAATATTCATT 273  
QY 81 GluAlaMetProThrPheLeuPheLeuLysAspGlyGluIleValAspLysValAlaGly 100  
|||||  
DB 274 GAGGCCATGCGCAACCTTCTCTTGAAGATGCGAGATCTGAGCAAGGTGTGTGT 333  
QY 101 AlaSerLysAspLysLeuGlnAlaThrIleAlaLysHisAlaSerAlaValAlaAla 120  
DB 334 GCTAGTAAGATGATCCTTCAAGCCACCATAGCCCAAGCATGATCTGCTGTGCTGCT 393  
QY 121 SerSerSer 123  
|||||  
DB 394 TCTTCTTCT 402

RESULT 4 BG238189 464 bp mRNA linear EST 28-NOV-2001  
LOCUS BG238189  
DEFINITION sab63d03.y1 Gm-c1043 glycine max cDNA clone GENOME SYSTEMS CLONE  
ID: Gm-c1043-3990 5' similar to SW:TH11\_TOBAC Q07090 THIOREDOXIN  
H-TYPE 2' mRNA sequence.

ACCESSION BG238189





fragments followed by NotI digestion. The cDNA fragments were directionally cloned into the NotI-SalI restriction site of the pSPORT1 vector. The ligated cDNA fragments were transformed into E. coli Electro-Max DH10B host cells. This library was constructed in the laboratory of Dr. Lila Vodkin by Anu Khanna at the University of Illinois at Urbana-Champaign. email: l-vodkin@uiuc.edu"

BASE COUNT 120 a 98 c 113 g 134 t

## ALIGNMENT SCORES:

Pred. No.: 3,56e-75 Length: 465  
Score: 618.00 Matches: 123  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 100.00% Indels: 0  
DB: 9 Gaps: 0

US-09-786-715-8 (1-123) x AM781A79 (1-465)

QY 1 MetAlaGluValGluGluGluValIleGlyValHisThrValAspLutrrpLysLeu 20  
DB 53 ATGGCTGAAGTGAAGAGGAGGACAGTCATCGCGCTCCACACCGCTTGATGAGGAAGCTG 112  
QY 21 GlnLeuGlnAsnAlaLysAspSerLysLysLeuIleValValAspPheThrAlaSerTrp 40  
DB 113 CAACCTCCAGAAATGCAAAAGACTCCAAAACATGATTTGCTGATTTACTCTCTCTG 172  
QY 41 CysGlyProCysArgPheMetAlaProValLeuAlaGluIleAlaLysLysThrProGlu 60  
DB 173 TGTGTGTCATGCGCGTTTATGCCCCAGCTTCTTGACAGATGCAAAAGAACTCTGTGA 232  
QY 61 LeuIlePheLeuLysValAspValAspGluValArgProValAlaGluLutrrpSerIle 80  
DB 233 TTGATCTTCCTCAAAAGTGGATGTGATGAGTGAAGCCCTGCTGAGCAATATTCATT 292  
QY 81 GluAlaMetProThrPheLeuPheLeuLysAspGlyGluIleValAlaAspLysValAlaGly 100  
DB 293 GAGGCCATGCGCAACCTCTCTTCTTGAAAGATGGAGATGTGGACAAAGTGTGTGT 352  
QY 101 AlaSerLysAspAspLeuGlnAlaThrIleAlaLysSHSAlaSerAlaValAlaAlaAla 120  
DB 353 GCTACTAAGAGATGACCTTCACACCAACATAGCAAGCATGATCTGCTGCTGCTGCT 412  
QY 121 SerSerSer 123  
DB 413 TCTTCTCTCT 421

RESULT 6 BE347100 475 bp mRNA linear EST 04-DEC-2001  
LOCUS BE347100 sp34h10.y1 Gm-cl043 Glycine max cDNA clone GENOME SYSTEMS CLONE ID:  
DEFINITION 2; mRNA sequence.  
2; mRNA sequence.

ACCESSION BE347100  
VERSION BE347100.1 GI:9258953  
KEYWORDS EST.  
SOURCE Glycine max (soybean)  
ORGANISM Glycine max

REFERENCE 1 (bases 1 to 475)  
AUTHORS Shoemaker R., Keim P., Vodkin L., Erpelting J., Coryell V., Khanna A., Bolla B., Maria M., Hillier L., Kuaba T., Matlin J., Beck C., Wylie T., Underwood K., Steptoe M., Theising B., Allen M., Bowers Y., Person B., Swaller T., Gibbons M., Pape D., Harvey N., Schurk R., Ritter E., Kohn S., Shin T., Jackson Y., Cardenas M., McCann R., Waterston R. and Wilson R.

TITLE Public Soybean EST Project  
JOURNAL Unpublished  
COMMENT Contact: Shoemaker R./Public Soybean EST Project

Public Soybean EST Project  
Washington University School of Medicine  
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA  
Tel: 314 286 1800  
Fax: 314 286 1810  
Email: est@wustl.wustl.edu  
This clone is available through: Resgen, Invitrogen Corp, 2130 South Memorial Parkway Huntsville, AL 35801 For further information call: (800)-533-4363 or contact via email: cou@resgen.com  
Insert Length: 906 Std Error: 0.00  
High quality sequence stop: 475.

## FEATURES

1..475  
Location/Qualifiers  
/organism="Glycine max"  
/mol\_type="mRNA"  
/db\_xref="taxon:3847"  
/clone="GENOME SYSTEMS CLONE ID: Gm-cl043-188"  
/tissue\_type="Hypocotyl and Plumule, germinating seeds"  
/lab\_host="DH10B"  
/clone\_1ib="Gm-cl043"  
/note="Vector: pTV73pac (Pharmacia); Site\_1: EcoRI; Site\_2: NotI; This cDNA library was constructed from mRNA isolated from hypocotyl and plumule tissues of seeds germinated for three days of the cultivar Williams. Complementary DNA was synthesized from mRNA using a primer consisting of a poly(dT) sequence with a NotI restriction site. EcoRI adapters were ligated to the blunt-ended cDNA fragments followed by digestion with EcoRI and NotI. The cDNA fragments were directionally cloned into the EcoRI-NotI restriction site of the pTV73-pac vector. The ligated cDNA fragments were transformed into DH10B host cells (Gibco BRL). This library was constructed by Dr. Randy Shoemaker."

BASE COUNT 126 a 97 c 118 g 134 t

## ALIGNMENT SCORES:

Pred. No.: 3.67e-75 Length: 475  
Score: 618.00 Matches: 123  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 100.00% Indels: 0  
DB: 10 Gaps: 0

US-09-786-715-8 (1-123) x BE347100 (1-475)

QY 1 MetAlaGluValGluGluGluValIleGlyValHisThrValAspLutrrpLysLeu 20  
DB 44 ATGGCTGAAGTGAAGAGGAGGACAGTCATCGCGCTCCACACCGCTTGATGAGGAAGCTG 103  
QY 21 GlnLeuGlnAsnAlaLysAspSerLysLysLeuIleValValAspPheThrAlaSerTrp 40  
DB 104 CAACCTCCAGAAATGCAAAAGACTCCAAAACATGATTTGCTGATTTACTCTCTCTG 163  
QY 41 CysGlyProCysArgPheMetAlaProValLeuAlaGluIleAlaLysLysThrProGlu 60  
DB 164 TGTGTGTCATGCGCGTTTATGCCCCAGCTTCTTGACAGATGCGCAAAAGAACTCTGAA 223  
QY 61 LeuIlePheLeuLysValAspValAspGluValArgProValAlaGluLutrrpSerIle 80  
DB 224 TTGATCTTCCTCAAAAGTGGATGTGATGAGTGAAGCCCTGCTGAGCAATATTCATT 283  
QY 81 GluAlaMetProThrPheLeuPheLeuLysAspGlyGluIleValAlaAspLysValAlaGly 100  
DB 284 GAGGCCATGCGCAACCTCTCTTCTTGAAAGATGGAGATGTGGACAAAGTGTGTGT 343  
QY 101 AlaSerLysAspAspLeuGlnAlaThrIleAlaLysSHSAlaSerAlaValAlaAlaAla 120  
DB 344 GCTACTAAGAGATGACCTTCACACCAACATAGCAAGCATGATCTGCTGCTGCTGCT 403  
QY 121 SerSerSer 123  
DB 404 TCTTCTCTCT 412

RESULT 7  
BE807019  
LOCUS  
DEFINITION  
479 bp mRNA linear EST 06-DEC-2001  
ss10e10.y1 Gm-cl047 glycine max cDNA clone GENOME SYSTEMS CLONE ID:  
Gm-cl047-2563 5' similar to SW:THIL\_TOBAC 007090 THIREDOXIN H-TYPE  
2 ; , mRNA sequence.  
ACCESSION  
BE807019  
VERSION  
KEYWORDS  
SOURCE  
ORGANISM  
glycine max (soybean)  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids  
; eurosids I; Fabales; Fabaceae; Papilionoideae; Phaseoleae;  
Glycine.  
REFERENCE  
1 (bases 1 to 479)  
AUTHORS  
Shoemaker, R., Keim, P., Vodkin, L., Erpelting, J., Coryell, V., Khanna  
, A., Bolla, B., Marra, M., Hillier, L., Kucaba, T., Martin, J., Beck, C.,  
Wylie, T., Underwood, K., Steptoe, M., Theising, B., Allen, M., Bowers  
, Y., Person, B., Swaller, T., Gibbons, M., Pape, D., Harvey, N., Schurk  
, R., Ritter, E., Kohn, S., Shin, T., Jackson, Y., Cardenas, M., McCann  
, R., Waterston, R. and Wilson, R.  
TITLE  
Public Soybean EST Project  
JOURNAL  
Unpublished  
COMMENT  
Contact: Shoemaker R./Public Soybean EST Project  
Public Soybean EST Project  
Washington University School of Medicine  
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA  
Tel: 314 286 1800  
Fax: 314 286 1810  
Email: est@watson.wustl.edu  
This clone is available through: Resgen, Invitrogen Corp. 2130  
South Memorial Parkway Huntsville, AL 35801 For further information  
call: (800)-533-4363 or contact via email: ccu@resgen.com  
High quality sequence stop: 449.  
Location/Qualifiers  
1. 479  
FEATURES  
source  
/organism="Glycine max"  
/mol\_type="mRNA"  
/db\_xref="taxon:3847"  
/clone="GENOME SYSTEMS CLONE ID: Gm-cl047-2563"  
/tissue\_type="Immature leaves (unfurled trifoliolate) of  
greenhouse grown plants"  
/dev\_stage="2 week old"  
/lab\_host="DH10B"  
/clone\_lib="Gm-cl047"  
/note="Vector: pSPORT1; Site1: NotI; Site2: SalI; This  
cDNA library was constructed from mRNA isolated from  
Immature leaves (unfurled trifoliolate) of greenhouse grown  
plants that were 2 week old. The library was prepared  
using the Life Technologies pSPURScript cDNA library  
construction kit. Complementary DNA was synthesized from  
mRNA using a poly(dT) sequence with a NotI restriction  
site. SalI linkers adapters were ligated to the  
blunt-ended cDNA fragments followed by NotI digestion. The  
cDNA fragments were directionally cloned into the  
NotI-SalI restriction site of the pSPORT1 vector. The  
ligated cDNA fragments were transformed into E.coli  
Electromax DH10B host cells. This library was constructed  
in the laboratory of Dr. Lila Vodkin by Anu Khanna at the  
University of Illinois at Urbana-Champaign (email  
l-vodkin@uiuc.edu)"

BASE COUNT  
ORIGIN  
128 a 88 c 122 g 141 t

Alignment Scores:  
Pred. No.: 3.71e-75  
Score: 618.00  
Percent Similarity: 100.00%  
Best Local Similarity: 100.00%  
Query Match: 100.00%  
DB: 10

Length: 479  
Matches: 123  
Conservative: 0  
Mismatch: 0  
Indels: 0  
Gaps: 0

us-09-786-715-8 (1-123) x BE807019 (1-479)  
QY  
1 MetLaGluValIGluGluValIleGlyValHisThrValAspGluTrpLysLeu 20  
Db  
18 ATGGCTGAAGTGAAGGAGGACAGCTTCGCGCTCCACACCGTTGATGATGGAAGCTG 77  
QY  
21 GlnLeuGlnAsnAlaLysAspSerLysLysLeuIleValValAspPheThrAlaSerTrp 40  
Db  
78 CAATCCAGATGCAAAAAGACTCAAAAACATGATTTGGTGGATTTTACGCTCCCG 137  
QY  
41 CysGlyProCysArGpHeMetAlaProValLeuAlaGluIleAlaLysLysThrProGlu 60  
Db  
138 TGTGCTCATGCGCTTTTATGCGCCCGCTTCAGAGATTCGCAAGAAACCTCGAA 197  
QY  
61 LeuIlePheLeuLysValAspValAspGluValAlaProValAlaGluIleTrpSerIle 80  
Db  
198 TTGATCTTCTCAAGATGGATGATGATGATGATGATGATGATGATGATGATGAT 257  
QY  
81 GluAlaMetProThrPheLeuPheLeuLysAspGlyLeuIleValAspLysValAlaGly 100  
Db  
258 GAGGCCATGCCACCTTCCTCTCTTGAAGATGCCAGATCGTGACAGGTGTTGCT 317  
QY  
101 AlaSerLysAspAspLeuGlnAlaThrIleAlaLysHisAlaSerAlaValAlaAla 120  
Db  
318 GCTAGTAAAGATGACCTTCAAGCCACCATAGCCAAAGATGATCTGCTGCTGCT 377  
QY  
121 SerSerSer 123  
Db  
378 TCTTCTTCT 386

RESULT 8  
BG725950  
LOCUS  
DEFINITION  
505 bp mRNA linear EST 29-NOV-2001  
sscl1004.y1 Gm-cl067 glycine max cDNA clone GENOME SYSTEMS CLONE  
ID: Gm-cl067-2383 5' similar to SW:THIL\_TOBAC 007090 THIREDOXIN  
H-TYPE 2 ; , mRNA sequence.  
ACCESSION  
BG725950  
VERSION  
KEYWORDS  
SOURCE  
ORGANISM  
glycine max (soybean)  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids  
; eurosids I; Fabales; Fabaceae; Papilionoideae; Phaseoleae;  
Glycine.  
REFERENCE  
1 (bases 1 to 505)  
AUTHORS  
Shoemaker, R., Keim, P., Vodkin, L., Erpelting, J., Coryell, V., Khanna  
, A., Bolla, B., Marra, M., Hillier, L., Kucaba, T., Martin, J., Beck, C.,  
Wylie, T., Underwood, K., Steptoe, M., Theising, B., Allen, M., Bowers  
, Y., Person, B., Swaller, T., Gibbons, M., Pape, D., Harvey, N., Schurk  
, R., Ritter, E., Kohn, S., Shin, T., Jackson, Y., Cardenas, M., McCann  
, R., Waterston, R. and Wilson, R.  
TITLE  
Public Soybean EST Project  
JOURNAL  
Unpublished  
COMMENT  
Contact: Shoemaker R./Public Soybean EST Project  
Public Soybean EST Project  
Washington University School of Medicine  
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA  
Tel: 314 286 1800  
Fax: 314 286 1810  
Email: est@watson.wustl.edu  
This clone is available through: Resgen, Invitrogen Corp. 2130  
South Memorial Parkway Huntsville, AL 35801 For further information  
call: (800)-533-4363 or contact via email: ccu@resgen.com  
High quality sequence stop: 419.  
Location/Qualifiers  
1. 505  
FEATURES  
source  
/organism="Glycine max"  
/mol\_type="mRNA"  
/db\_xref="taxon:3847"  
/tissue\_type="Germinating shoot, 3 day old seedling, auxin

treatment"  
/lab\_host="DH10B"  
/clone\_lib="Gm-cl067"  
/note="Vector: pBluescript II SK-; Site.1: EcoRI, Site.2: XhoI; The cDNA library was constructed from mRNA isolated from germinating shoots of 3 day old seedling by XhoI cDNA library Williams 82. The seedlings were germinated in a growth chamber using germination paper in a solution containing 100ppm auxin. Complementary DNA was synthesized from mRNA using a primer consisting of a poly(dT) sequence with a XhoI restriction site. EcoRI adapters were ligated to the blunt-ended cDNA fragments followed by XhoI digestion. The cDNA fragments were directionally cloned into the EcoRI-XhoI restriction site of the pBluescript vector. The ligated cDNA fragments were transformed into DH10B host cells (GibcoBRL). This library was constructed in the laboratory of Dr. Randy Shoemaker."

BASE COUNT 141 a 91 c 128 g 145 t

ORIGIN

Alignment Scores:  
Pred. No.: 3,99e-75 Length: 505  
Score: 618.00 Matches: 123  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 100.00% Indels: 0  
DB: 10 Gaps: 0

US-09-786-715-8 (1-123) x BG725950 (1-505)

QY 1 MetAlaGluValGluGluGluGluValIleGlyValHisThrValAspGluTrpLysLeu 20  
DB 11 ATGGCTGAAGTGGAGAGGAGACAGTCATCGGCTCCACCCCTGATGATGAGAGCTG 70  
QY 21 GlnLeuGlnAsnAlaLysAspSerLysLysLeuIleValValAspPheThrAlaSerTrp 40  
DB 71 CAACTCCAGATGCCAAAGACTCCAAAACATGATGATGATGATGATGATGATGATGATG 130  
QY 41 CysGlyProCysArgPheMetAlaProValLeuAlaGluIleAlaLysLysThrProGlu 60  
DB 131 TGTGGTCCATGCCGCTTTATGCCCCAGTTCGACAGATTCGCAAGAAACCTCCTGAA 190  
QY 61 LeuIlePheLeuLysValAspValAspGluValArgProValAlaGluGluTrpSerIle 80  
DB 191 TTGATCTTCTCAAAATGATGATGATGATGATGATGATGATGATGATGATGATGATG 250  
QY 81 GluAlaMetProThrPheLeuPheLeuLysAspGlyGluIleValAspLysValAlaGly 100  
DB 251 GAGGCCATGCAACCTCTCTCTCTGAAAGATGCGAGATCGTGGAAGGTGGTGGT 310  
QY 101 AlaSerLysAspLysLeuGlnAlaThrIleAlaLysHisAlaSerAlaValAlaAla 120  
DB 311 GCTAGTAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 370  
QY 121 SerSerSer 123  
DB 371 TCTTCTTCT 379

RESULT 9  
BE607957 508 bp mRNA linear EST 04-DEC-2001  
LOCUS BE607957  
DEFINITION sq19B02.y1 Gm-cl046 glycine max cDNA clone GENOME SYSTEMS CLONE ID: Gm-cl046-1252.5' similar to SW:TH11\_TOBAC Q07090 THIOREDOXIN H-TYPE 2; mRNA sequence.

ACCESSION BE607957  
VERSION BE607957.1 GI:9898989  
KEYWORDS EST.  
SOURCE glycine max (soybean)  
ORGANISM Glycine max  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids; eurosids I; Fabales; Fabaceae; Papilionoideae; Phaseoleae; Glycine.

REFERENCE 1 (bases 1 to 508)  
AUTHORS Shoemaker, R., Keim, P., Vodkin, L., Erpellding, J., Corvett, V., Khanna, A., Bolla, B., Marra, M., Hillier, L., Kucaba, T., Martin, J., Beck, C., Wylie, T., Underwood, K., Steptoe, M., Treisman, B., Allen, M., Bowers, Y., Person, B., Swaller, T., Gibbons, M., Page, D., Harvey, N., Schurk, R., Ritter, E., Kohn, S., Shin, T., Jackson, Y., Cardenas, M., McCann, R., Waterston, R. and Wilson, R.  
TITLE Public Soybean EST Project  
JOURNAL Unpublished  
COMMENT Contact: Shoemaker R/Public Soybean EST Project  
Public Soybean EST Project  
Washington University School of Medicine  
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA  
Tel: 314 286 1800  
Fax: 314 286 1810  
Email: est@wustl.edu  
This clone is available through: Resgen, Invitrogen Corp. 2130 South Memorial Parkway Huntville, AL 35801 For further information call: (800)-533-4363 or contact via email: cou@resgen.com  
Insert length: 731 Std Error: 0.00  
High quality sequence stop: 476.  
Location/Qualifiers  
1. 508  
/organism="Glycine max"  
/mol\_type="mRNA"  
/db\_xref="taxon:3847"  
/clone="GENOME SYSTEMS CLONE ID: Gm-cl046-1252"  
/tissue="Germinating Seeds"  
/lab\_host="DH10B"  
/clone\_lib="Gm-cl046"  
/note="Vector: PT73Pac (Pharmacia); Site.1: EcoRI; Site.2: NotI; This cDNA library was constructed from mRNA isolated from seeds germinated for three days of the cultivar Williams. Complementary DNA was synthesized from mRNA using a primer consisting of a poly(dT) sequence with a NotI restriction site. EcoRI adapters were ligated to the blunt-ended cDNA fragments followed by digestions with EcoRI and NotI. The cDNA fragments were directionally cloned into the EcoRI-NotI restriction site of the PT73-Pac vector. The ligated cDNA fragments were transformed into DH10B host cells (Gibco BRL). This library was constructed by Dr. Randy Shoemaker."

BASE COUNT 140 a 93 c 128 g 147 t

ORIGIN

Alignment Scores:  
Pred. No.: 4.02e-75 Length: 508  
Score: 618.00 Matches: 123  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 100.00% Indels: 0  
DB: 10 Gaps: 0

US-09-786-715-8 (1-123) x BE607957 (1-508)

QY 1 MetAlaGluValGluGluGluGluValIleGlyValHisThrValAspGluTrpLysLeu 20  
DB 1 ATGGCTGAAGTGGAGAGGAGACAGTCATCGGCTCCACCCCTGATGATGAGAGCTG 60  
QY 21 GlnLeuGlnAsnAlaLysAspSerLysLysLeuIleValValAspPheThrAlaSerTrp 40  
DB 61 CAACTCCAGATGCCAAAGACTCCAAAACATGATGATGATGATGATGATGATGATGATG 120  
QY 41 CysGlyProCysArgPheMetAlaProValLeuAlaGluIleAlaLysLysThrProGlu 60  
DB 121 TGTGGTCCATGCCGCTTTATGCCCCAGTTCGACAGATGCAAGAAACCTCCTGAA 180  
QY 61 LeuIlePheLeuLysValAspValAspGluValArgProValAlaGluGluTrpSerIle 80  
DB 181 TTGATCTTCTCAAAATGATGATGATGATGATGATGATGATGATGATGATGATGATG 240  
QY 81 GluAlaMetProThrPheLeuPheLeuLysAspGlyGluIleValAspLysValAlaGly 100

Db 241 GAGGCATGACCACTTCTCTCTTGAAGATGCGAGATCGTGACAGGTGCTGT 300  
 QY 101 AlaserlyasbapleuGlnAlaThrlleAlaLysHsAlaserAlaValAlaAla 120  
 Db 301 GCTACTAAGATGACTTCAAGCCACATAGCAGATGATGCTGTGCTGTCT 360  
 QY 121 SerSerSer 123  
 Db 361 TCTCTCTCT 369  
 RESULT 10  
 AM568753  
 LOCUS  
 DEFINITION  
 510 bp mRNA linear EST 03-DEC-2001  
 s172f07.y1 Gm-c1031 Glycine max cDNA clone GENOME SYSTEMS CLONE ID:  
 Gm-c1031-254 5' similar to SW:TH11\_TOBAC Q07090 THIOREDOXIN H-TYPE  
 2 ; mRNA sequence.  
 AM568753  
 ACCESSION  
 VERSION  
 KEYWORDS  
 SOURCE  
 ORGANISM  
 Glycine max (soybean)  
 EST.  
 Glycine max  
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids  
 ; eurosids I; Fabales; Fabaceae; Papilionoideae; Phaseoleae;  
 Glycine.  
 1 (bases 1 to 510)  
 Shoemaker,R., Keim,P., Vodkin,L., Erpelting,J., Corryell,V., Khanna  
 ,A., Bolla,B., Marra,M., Hillier,L., Kucaba,T., Martin,J., Beck,C.,  
 Wylie,T., Underwood,K., Steptoe,M., Theising,B., Allen,M., Bowers  
 ,I., Ferson,B., Swaller,T., Gibbons,M., Pape,D., Harvey,N., Schurk  
 ,R., Rittler,E., Kohn,S., Shih,T., Jackson,Y., Cardenas,M., McCann  
 ,R., Waterston,R. and Wilson,R.  
 Public Soybean EST Project  
 Unpublished  
 Contact: Shoemaker R./Public Soybean EST Project  
 Public Soybean EST Project  
 Washington University School of Medicine  
 4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA  
 Tel: 314 286 1800  
 Fax: 314 286 1810  
 Email: est@watson.wustl.edu  
 This clone is available through: Resgen, Invitrogen Corp. 2130  
 South Memorial Parkway Huntsville, AL 35801 For further information  
 call: (800)-533-4363 or contact via email: cdueresgen.com  
 Insert Length: 729 Std Error: 0.00  
 High quality sequence stop: 411.  
 Location/Qualifiers  
 1..510  
 /organism="Glycine max"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:3847"  
 /clone="GENOME SYSTEMS CLONE ID: Gm-c1031-254"  
 /tissue\_type="Williams seedlings, minus the cotyledons"  
 /lab\_host="DH10B"  
 /clone\_lib="Gm-c1031"  
 /note="Vector: pBluescript II XR; Site1: EcoRI; Site2:  
 XhoI; This cDNA library was constructed from mRNA isolated  
 from whole 'Williams' seedlings, minus the cotyledons,  
 which were propagated on paper towels with distilled water  
 for 5 days. Incubated at 40 degrees C for 1 hour. The  
 cotyledons were removed and the remaining tissue was  
 flash-frozen in liquid nitrogen. Stratagene's cDNA  
 Synthesis Kit (catalog number 200401) was used to  
 synthesize the cDNA. First-strand synthesis was performed  
 with 5-methyl dCTP, hence the ligated cDNA was  
 hemimethylated. A modification of Stratagene's  
 first-strand synthesis primer was used. An 'anchor'  
 nucleotide (V-A, C, or G) was added to the 3' end of the  
 primer [GAGAGAGAGAGAGAGAGAGAGACTGAG(T)18V] to anchor  
 the primer at the 5' end of the poly(A) tract. After  
 second strand synthesis, the cDNA ends were filled in with  
 cloned Pfu DNA, ligated to EcoRI adapters and  
 subsequently phosphorylated. The cDNA was then

precipitated and redissolved in sterile, RNase-,  
 DNase-free water. The XhoI site within the first-strand  
 synthesis primer was then restricted by digestion with  
 XhoI from Promega (400U/ul); all XhoI sites in the cDNA  
 would be protected by their hemimethylated status. The  
 cDNA constructs were size-fractionated with a 500 bp  
 cutoff, using GIBCOBRL Life Technologies' cDNA Size  
 Fractionation column. The column eluent was then  
 precipitated, redissolved, and ligated into Stratagene's  
 pBluescript II XR predigested vector (pBluescript II SK(+)  
 vector that has been digested with EcoRI and XhoI, and  
 phosphorylated by Stratagene). 100% of the white and blue  
 colonies appear to contain recombinant plasmids with cDNA  
 inserts, based on size (n=18 and 5, respectively). This  
 library was constructed by Dr. Paul Keim and Dr. Virginia  
 Corryell."

BASE COUNT 142 a 92 c 129 g 147 t  
 ORIGIN  
 Alignment Scores:  
 Pred. No.: 4,04e-75 Length: 510  
 Score: 618.00 Matches: 123  
 Percent Similarity: 100.00% Conservative: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 100.00% Indels: 0  
 DB: 9 Gaps: 0  
 US-09-786-715-8 (1-123) x AM568753 (1-510)  
 QY 1 MetAlaGlnValGlnGluGlyGlnValIleGlyValAlaHsThrValAspGluTrpLysLeu 20  
 Db 18 ATGCTGAAGTGAAGGAGGACAGTCATCCGCTCCACACCGCTTGATGATGCAACCTG 77  
 QY 21 GlnLeuGlnAsnAlaLysAspSerLysLysLeuIleValAlaAspPheThrAlaSerTrp 40  
 Db 78 CAATCCAGATGCAAAAAGACTCCAAAAGACTGATGGTGAGATTACGCTTCTCTCG 137  
 QY 41 CysGlyProCysArgPheMetAlaProValLeuAlaGluIleAlaLysLysThrProGlu 60  
 Db 138 TGTGTCATCCATCCGCTTTATGCCCCAGTCTTCAGAGATTCGAAGAAACCTCGAA 197  
 QY 61 LeuIlePheLeuLysValAspValAspGlnValArgProValAlaGluTrpSerIle 80  
 Db 198 TTGATCTTCCTCAAGATGATGATGATGAAGTGAAGCTGTGCTGAGGAATATTCCATT 257  
 QY 81 GlnAlaMetProThrPheLeuPheLeuLysAspGlyGluIleValAlaAspLysValIleGly 100  
 Db 258 GAGCCATGCAACCTTCTCTCTTGAAGATGCGCAGATCGCGACAAAGTGTGCT 317  
 QY 101 AlaserlyasbapleuGlnAlaThrlleAlaLysHsAlaserAlaValAlaAla 120  
 Db 318 GCTAGTAAAGATGACTTCAAGCCACATAGCAGATGATGCTGTGCTGTCT 377  
 QY 121 SerSerSer 123  
 Db 378 TCTCTCTCT 386  
 RESULT 11  
 AM164347  
 LOCUS  
 DEFINITION  
 517 bp mRNA linear EST 02-DEC-2001  
 se1f11.y1 Gm-c1023 Glycine max cDNA clone GENOME SYSTEMS CLONE ID:  
 Gm-c1023-141 5' similar to SW:TH11\_TOBAC Q07090 THIOREDOXIN H-TYPE  
 2 ; mRNA sequence.  
 AM164347  
 ACCESSION  
 VERSION  
 KEYWORDS  
 SOURCE  
 ORGANISM  
 Glycine max (soybean)  
 EST.  
 Glycine max  
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids  
 ; eurosids I; Fabales; Fabaceae; Papilionoideae; Phaseoleae;  
 Glycine.  
 1 (bases 1 to 517)

## AUTHORS

Shoemaker, R., Kelm, P., Vodkin, L., Erpelding, J., Corryell, V., Khanna, A., Bolla, B., Marra, M., Hillier, L., Kucaba, T., Martin, J., Beck, C., Wylie, T., Underwood, K., Steptoe, M., Theising, B., Allen, M., Bowers, Y., Person, B., Smaller, T., Gibbons, M., Pape, D., Harvey, N., Schurk, R., Ritter, E., Kohn, S., Shin, T., Jackson, Y., Cardenas, M., McCann, R., Waterston, R. and Wilson, R.

## TITLE

Public Soybean EST Project

## JOURNAL

Unpublished

## COMMENT

Contact: Shoemaker R/Public Soybean EST Project  
Public Soybean EST Project  
Washington University School of Medicine  
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA  
Tel: 314 286 1800  
Fax: 314 286 1810  
Email: est@watson.wustl.edu  
This clone is available through: Resgen, Invitrogen Corp. 2130  
South Memorial Parkway Huntville, AL 35601 For further information  
call: (800)-533-4363 or contact via email: coueresgen.com  
Insert Length: 782 Std Error: 0.00  
High quality sequence stop: 417.

## FEATURES

Location/Qualifiers

1..517  
/organism="Glycine max"  
/mol\_type="mRNA"  
/db\_xref="taxon:3847"  
/clone="GENOME SYSTEMS CLONE ID: Gm-cl023-141"  
/tissue\_type="seed coats of greenhouse grown plants"  
/lab\_host="DH10B"  
/clone\_lib="Gm-cl023"  
/note="Vector: pSPORT1; Site-1: SalI; Site-2: NotI; This  
cDNA library was constructed from mRNA isolated from seed  
coats (100-200mg) of greenhouse grown plants. The  
library was prepared using the Life Technologies  
SuperScript cDNA library construction kit. Complementary  
DNA was synthesized from mRNA using a poly (dT) sequence  
with a Not I restriction site. Sal I linkers adapters  
were ligated to the blunt-ended cDNA fragments followed by  
Not I digestion. The cDNA fragments were directionally  
cloned into the Not I-Sal I restriction site of the  
pSPORT1 vector. The ligated cDNA fragments were  
transformed into E.coli Electromax DH10B host cells. This  
library was constructed by Dr. Lila Vodkin and Dr. Anu  
Khanna."

BASE COUNT 140 a 96 c 129 g 152 t  
ORIGIN

## Alignment Scores:

Pred. No.: 4.12e-75 Length: 517  
Score: 618.00 Matches: 123  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 100.00% Indels: 0  
DB: 9 Gaps: 0

US-09-786-715-8 (1-123) x AWI64347 (1-517)

QY 1 Metatagluvalglugluclyglvalilleglyvalhsthvvalasplutrrplysleu 20  
DB 27 ATGGCTGAAGTGAAGAGGACAGGACAGGCGGCGCCACACCGCTGATGAGGAAGCTG 86  
QY 21 Glnleuglnasnalylaspserylsylsleuilevalvalaspphenthralasertp 40  
DB 87 CAACCTCCAGAAATGCAAAAGACTCCAAAACACTGATTTGGTGATTTTACTGCTTCTG 146  
QY 41 Cysglyprocyasarphehmetalarovalleualagluilealalyslysthrproglu 60  
DB 147 TGTGTTCATGCGCTTTATGCGCCAGTTCTTGCAAGATTCGAAAGAACTCCTGAA 206  
QY 61 Leuilepheleulyvalasplvalasplvalasplvalasplvalasplvalaspl 80  
DB 207 TTGATCTTCTCCTCAAGTGGATGTGATGAAGGAGCCGCTGAGGAAATATTCATT 266  
QY 81 Gualametprothrphleupheleulyaspsglygluilevalasplvalvalgily 100

## DB

|||||  
267 GAGGCATGCCAACCTCTCTTCTGAAAGATGGAGATGCTGGACAAAGTGCTGCT 326

## QY

101 Alaserylaspaspplenglinalarhrilealalyslshlaserlavalalalala 120

## DB

327 GCTATGAGATGATGACCTTCACAGCACCATAGCAACGATCTGCTGCTGCTGCT 386

## QY

121 SerSerSer 123  
|||||

## DB

387 TCTTCTTCT 395

## RESULT 12

BM884981 527 bp mRNA linear EST 08-MAR-2002  
LOCUS  
DEFINITION  
sal93d11.y1 Gm-cl063 Glycine max cDNA clone SOYBEAN CLONE ID:  
Gm-cl063-3622 5' similar to SW:TH1170BAC 007090 THIRODOXIN H-TYPE  
2', mRNA sequence.

## ACCESSION

BM884981  
BM884981.1 GI:19268725

## VERSION

EST.

## KEYWORDS

Glycine max (soybean)

## SOURCE

Glycine max (soybean)  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids  
; eurosids I; Fabales; Fabaceae; Papilionoideae; Phaseoleae;  
Glycine.

## REFERENCE

1 (bases 1 to 527)

## AUTHORS

Shoemaker, R., Kelm, P., Vodkin, L., Erpelding, J., Corryell, V., Khanna, A., Bolla, B., Marra, M., Hillier, L., Kucaba, T., Martin, J., Beck, C., Wylie, T., Underwood, K., Steptoe, M., Theising, B., Allen, M., Bowers, Y., Person, B., Smaller, T., Gibbons, M., Pape, D., Harvey, N., Schurk, R., Ritter, E., Kohn, S., Shin, T., Jackson, Y., Cardenas, M., McCann, R., Waterston, R. and Wilson, R.

## TITLE

Public Soybean EST Project

## JOURNAL

Unpublished

## COMMENT

Contact: Shoemaker R/Public Soybean EST Project  
Public Soybean EST Project  
Washington University School of Medicine  
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA  
Tel: 314 286 1800  
Fax: 314 286 1810  
Email: est@watson.wustl.edu  
This clone is available through: Resgen, Invitrogen Corp. 2130  
South Memorial Parkway Huntville, AL 35601 For further information  
call: (800)-533-4363 or contact: coueresgen.com web site:  
www.resgen.com  
Seq primer: -40RP from Gibco  
High quality sequence stop: 421.

## FEATURES

Location/Qualifiers

1..527  
/organism="Glycine max"  
/mol\_type="mRNA"  
/db\_xref="taxon:3847"  
/clone="SOYBEAN CLONE ID: Gm-cl063-3622"  
/tissue\_type="germinating shoot, 24 hour germination"  
/lab\_host="DH10B"  
/clone\_lib="Gm-cl063"  
/note="Vector: pBluescript II SK+; Site-1: EcoRI; Site-2:  
XhoI; The cDNA library was constructed from mRNA isolated  
from germinating shoots of the cultivar Williams. The  
seeds were allowed to germinate for 24 hours prior to  
harvesting the germinating shoots. Complementary DNA was  
synthesized from mRNA using a primer consisting of a  
poly(dT) sequence with a XhoI restriction site. EcoRI  
adapters were ligated to the blunt-ended cDNA fragments  
followed by XhoI digestion. The cDNA fragments were  
directionally cloned into the EcoRI-XhoI restriction site  
of the pBluescript vector. The ligated cDNA fragments were  
transformed into DH10B host cells (Gibco BRL). This  
library was constructed in the laboratory of Dr. Randy  
Shoemaker at Iowa state university."

BASE COUNT 144 a 98 c 130 g 155 t  
ORIGIN

## Alignment Scores:

Pred. No.: 4,236-75 Length: 527  
Score: 618.00 Matches: 123  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 100.00% Indels: 0  
DB: 12 Gaps: 0

US-09-786-715-8 (1-123) x BM884981 (1-527)

QY 1 MetlaagiuaValgluGluGluValIleGlyValIleHsrHrValAspGluTrpLysLeu 20  
DB 32 ATGGCTGAAGTGAAGAGGAGGACAGGTCAATCGCGCCACACCGTTGATGAGTGAAGCTG 91  
QY 21 GlnleuglnAsnAlaLysAspSerLysLysLeuIleValAspPheThrAlaSerTrp 40  
DB 92 CAATCCGAGATGCAAAAGACTCCAAAAGACTGATGGTGGATTTACTGCTTCCG 151  
QY 41 CysglProCysArgPheMetAlaProValLeuAlaGluIleAlaLysLysThrProGlu 60  
DB 152 TGTGGTCATCGCCGTTTATGAGCCAGTCTTGCAGAGATTGCAAGAAAACCTCCGAA 211  
QY 61 LeuIlePheLeuLysValAspValAspGluValArgProValAlaGluGluTrpSerIle 80  
DB 212 TTGATCTTCCTCAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGT 271  
QY 81 GluAlaMetProThrPheLeuPheLeuLysAspGlyGluIleValAspLysValAlaGly 100  
DB 272 GAGGCCATGCCACACTTCTCTTCTTGAAAGATGCCAGATCGTGACAAAGTGGTTGCT 331  
QY 101 AlaSerLysAspAspLeuGlnAlaThrIleAlaLysHisAlaSerAlaValAlaAla 120  
DB 332 GCTGTAAAGATGACCTTCAAGCCACCATACCAAGCATGATGCTGTGCTGCTGCT 391  
QY 121 SerSerSer 123  
DB 392 TCTTCTTCT 400

## RESULT 13

CD414498 532 bp mRNA linear EST 07-JUN-2003  
LOCUS Gm.c46654 Soybean induced by Salicylic Acid glycine max cDNA 3',  
DEFINITION mRNA sequence.

ACCESSION CD414498

VERSION CD414498.1 GI:31472470

KEYWORDS EST.

SOURCE Glycine max (soybean)

ORGANISM Glycine max (soybean)

Eudicotyledons; Streptophyta; Embryophyta; Tracheophyta;  
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids  
; eurosids I; Fabales; Fabaceae; Papilionoideae; Phaseoleae;  
Glycine.

## REFERENCE 1 (bases 1 to 532)

Tian A.-G., Wang J., Cui P., Han Y.-J., Xu H., Cong L.-J., Huang  
, X.-G., Wang X.-L., Jiao Y.-Z., Wang B.-J., Wang Y.-J., Zhang J.-S.  
, Chen S.-Y. and Yu J.  
Soybean Expressed Sequence Tags sequencing

Unpublished  
Contact: Chen S.-Y.  
Plant Biotechnology Laboratory  
Institute of Genetics and Developmental Biology, CAS, China  
Detun road, Beijing 100101, China  
Tel: 86-10-6486859  
Fax: 86-10-64873428  
Email: sychen@genetics.ac.cn  
Email: sychen@genetics.ac.cn  
Seq primer: 17 primer.

## TITLE

## JOURNAL

## COMMENT

## FEATURES

## source

1..532  
location/Qualifiers  
/organism="Glycine max"  
/mol\_type="mRNA"  
/cultur="Kefeng 1"

/db\_xref="taxon:3847"  
/tissue\_type="Seedlings"  
/dev\_stage="two-week seedlings"  
/lab\_host="Xil-Blue MRF strain"  
/clone\_lib="Soybean induced by Salicylic Acid"  
/note="Vector: pBluescript SK+; Site\_1: EcoR I; Site\_2:  
Xho I; The cDNA library was constructed by He, C-Y from  
mRNA isolated from two-week seedlings (cultivar Kefeng 1)  
treated by spraying 2.0mM salicylic acid for 24, 36, 48  
and 72 h. Complementary DNA was synthesized from mRNA  
using a primer consisting of a poly(dT) sequence with a  
XhoI restriction site. EcoRI adapters were ligated to the  
blunt-ended cDNA fragments followed by XhoI digestion. The  
cDNA fragments were directionally cloned into the  
EcoRI-XhoI restriction site of the pBluescript vector. The  
ligated cDNA fragments were transformed into Xil-Blue MRF  
host cells (Stratagene)."   
BASE COUNT 158 a 135 c 94 g 145 t  
ORIGIN

## Alignment Scores:

Pred. No.: 4,286-75 Length: 532  
Score: 618.00 Matches: 123  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 100.00% Indels: 0  
DB: 14 Gaps: 0

US-09-786-715-8 (1-123) x CD414498 (1-532)

QY 1 MetlaagiuaValgluGluGluValIleGlyValIleHsrHrValAspGluTrpLysLeu 20  
DB 511 ATGGCTGAAGTGAAGAGGAGGACAGTTCGCGCCACACCGTTGATGAGTGAAGCTG 452  
QY 21 GlnleuglnAsnAlaLysAspSerLysLysLeuIleValAspPheThrAlaSerTrp 40  
DB 451 CAATCCGAGATGCAAAAGACTCCAAAAGACTGATGGTGGATTTACTGCTTCCG 392  
QY 41 CysglProCysArgPheMetAlaProValLeuAlaGluIleAlaLysLysThrProGlu 60  
DB 391 TGTGGTCATCGCCGTTTATGAGCCAGTCTTGCAGAGATTGCAAGAAAACCTCCGAA 332  
QY 61 LeuIlePheLeuLysValAspValAspGluValArgProValAlaGluGluTrpSerIle 80  
DB 331 TTGATCTTCCTCAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGT 272  
QY 81 GluAlaMetProThrPheLeuPheLeuLysAspGlyGluIleValAspLysValAlaGly 100  
DB 271 GAGGCCATGCCACACTTCTCTTCTTGAAAGATGCCAGATCGTGACAAAGTGGTTGCT 212  
QY 101 AlaSerLysAspAspLeuGlnAlaThrIleAlaLysHisAlaSerAlaValAlaAla 120  
DB 211 GCTGTAAAGATGACCTTCAAGCCACCATACCAAGCATGATGCTGTGCTGCTGCT 152  
QY 121 SerSerSer 123  
DB 151 TCTTCTTCT 143

## RESULT 14

BE440397 544 bp mRNA linear EST 04-DEC-2001  
LOCUS sp45405.y1 Gm-c1043 Glycine max cDNA clone GENOME SYSTEMS CLONE ID:  
DEFINITION Gm-c1043-1161 5' similar to SW:TH11.TOBAC Q07090 THIOREDOXIN H-TYPE  
2; ; mRNA sequence.

ACCESSION BE440397

VERSION BE440397.1 GI:9439885

KEYWORDS EST.

SOURCE Glycine max (soybean)

ORGANISM Glycine max (soybean)

Eudicotyledons; Streptophyta; Embryophyta; Tracheophyta;  
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids  
; eurosids I; Fabales; Fabaceae; Papilionoideae; Phaseoleae;  
Glycine.

REFERENCE 1 (bases 1 to 544)  
 AUTHORS Shoemaker, R., Keim, P., Vodkin, L., Erpellding, J., Coryell, V., Khanna, A., Bolla, B., Marra, M., Hillier, L., Kucaba, T., Martin, J., Beck, C., Wylie, T., Underwood, K., Steptoe, M., Theising, B., Allen, M., Bowers, Y., Person, B., Swaller, T., Gibbons, M., Page, D., Harvey, N., Schurk, R., Ritter, E., Kohn, S., Shin, T., Jackson, Y., Cardenas, M., McCann, R., Waterston, R. and Wilson, R.  
 TITLE Public Soybean EST Project  
 JOURNAL Unpublished  
 COMMENT Contact: Shoemaker R/Public Soybean EST Project  
 Public Soybean EST Project  
 Washington University School of Medicine  
 4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA  
 Tel: 314 286 1800  
 Fax: 314 286 1810  
 Email: est@watson.wustl.edu  
 This clone is available through: Resgen, Invitrogen Corp. 2130 South Memorial Parkway Huntville, AL 35801 For further information call: (800)-533-4363 or contact via email: cu@resgen.com  
 Insert Length: 661 Std Error: 0.00  
 High quality sequence stop: 479.

FEATURES  
 source  
 1..544  
 /organism="Glycine max"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:3847"  
 /clone="GENOME SYSTEMS CLONE ID: Gm-c1043-1161"  
 /tissue\_type="Hypocotyl and Plumule, germinating seeds"  
 /lab\_host="DH10B"  
 /clone\_lib="Gm-c1043"  
 /note="Vector: pT73Pac (Pharmacia); Site.1: EcoRI; Site.2: NotI; This cDNA library was constructed from mRNA isolated from hypocotyl and plumule tissues of seeds germinated for three days of the cultivar Williams. Complementary DNA was synthesized from mRNA using a primer consisting of a poly(dT) sequence with a NotI restriction site. EcoRI adapters were ligated to the blunt-ended cDNA fragments followed by digestion with EcoRI and NotI. The cDNA fragments were directionally cloned into the EcoRI-NotI restriction site of the pT73-Pac vector. The ligated cDNA fragments were transformed into DH10B host cells (Gibco BRL). This library was constructed by Dr. Randy Shoemaker."

BASE COUNT 148 a 102 c 134 g 160 t  
 ORIGIN  
 Alignment Scores:  
 Pred. No.: 4.41e-75 Length: 544  
 Score: 618.00 Matches: 123  
 Percent Similarity: 100.00% Conservative: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 100.00% Indels: 0  
 DB: 10 Gaps: 0

US-09-786-715-8 (1-123) x BE440397 (1-544)

OY 1 MetAlaGluValGluGluGluGluValHisThrValAspGluTrpTyrSer 20  
 |||||  
 DB 44 ATGGCTGAAGTGGAAAGGACAGGTCATCGGCTCACCCGTTGATGAGTGAAGCTG 103  
 |||||  
 OY 21 GlnLeuGlnAsnAlaLysAspSerLysLysLeuIleValValAspPheThrAlaSerTrp 40  
 |||||  
 DB 104 CAACCTCCAGATGCAAAAGACTCCAAAAAAGTATGCTGTGATTTTACTGCTTCTG 163  
 |||||  
 OY 41 CysGlyProCysArgPheMetAlaProValLeuAlaGluIleAlaLysLysThrProGlu 60  
 |||||  
 DB 164 TGGGGTCATGCGCTTTATGGCCAGTCTTGACAGATGCAAAAGAAACTCCTGAA 223  
 |||||  
 OY 61 LeuIlePheLeuLysValAspValAspGluValArgProValAlaGluLysTrpSerIle 80  
 |||||  
 DB 224 TTGATTTCTCTCAAAATGAGTGTGATGAGGAGCGCTGCTCAGGAATATTTCCATT 283  
 |||||  
 OY 81 GluAlaMetProThrPheLeuPheLeuLysAspGlyGluIleValAspLysValValGly 100

DB 284 GAGGCATGCCACCTTCCCTCTTGAAGAAGCGCAGATCTCGACAGAGCTGGT 343  
 |||||  
 OY 101 AlASerLysAspAspLeuGlnAlaThrIleAlaLysHisAlaSerAlaValAlaAla 120  
 |||||  
 DB 344 GCTATGAAAGATACCTTCACAGCCACCATGACATGCATCTGCTGCTGCTGCT 403  
 |||||  
 OY 121 SerSerSer 123  
 |||||  
 DB 404 TCTTCTTCT 412

RESULT 15  
 BM177477  
 LOCUS  
 DEFINITION  
 569 bp mRNA linear EST 06-DEC-2001  
 Gm-c1074.3024 5' similar to SW:11111-TOBAC 007090 THIOREDOXIN H-TYPE 2 ; mRNA sequence.  
 BM177477  
 BM177477.1 GI:17400695  
 EST.  
 Glycine max (soybean)  
 Glycine max  
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids ; eurosids I; Fabales; Fabaceae; Papilionoideae; Phaseoleae; Glycine.  
 1 (bases 1 to 569)  
 REFERENCE  
 AUTHORS Shoemaker, R., Keim, P., Vodkin, L., Erpellding, J., Coryell, V., Khanna, A., Bolla, B., Marra, M., Hillier, L., Kucaba, T., Martin, J., Beck, C., Wylie, T., Underwood, K., Steptoe, M., Theising, B., Allen, M., Bowers, Y., Person, B., Swaller, T., Gibbons, M., Page, D., Harvey, N., Schurk, R., Ritter, E., Kohn, S., Shin, T., Jackson, Y., Cardenas, M., McCann, R., Waterston, R. and Wilson, R.  
 TITLE Public Soybean EST Project  
 JOURNAL Unpublished  
 COMMENT Contact: Shoemaker R/Public Soybean EST Project  
 Public Soybean EST Project  
 Washington University School of Medicine  
 4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA  
 Tel: 314 286 1800  
 Fax: 314 286 1810  
 Email: est@watson.wustl.edu  
 This clone is available through: Resgen, Invitrogen Corp. 2130 South Memorial Parkway Huntville, AL 35801 For further information call: (800)-533-4363 or contact: cu@resgen.com web site: www.resgen.com  
 Seq primer: -40RP from Gibco  
 High quality sequence stop: 437.

FEATURES  
 source  
 1..569  
 /organism="Glycine max"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:3847"  
 /clone="SOYBEAN CLONE ID: Gm-c1074-3024"  
 /tissue\_type="seedlings induced for HR (hypersensitive response)"  
 /dev\_stage="9-11 day old"  
 /lab\_host="DH10B"  
 /clone\_lib="Gm-c1074"  
 /note="Vector: pBluescript II SK+; Site.1: EcoRI; Site.2: XhoI; The cDNA library was constructed from mRNA isolated from 9-11 day old seedlings that were induced for HR (hypersensitive response) by vacuum infiltrating plant tissue with Pseudomonas syringae pv. glycinea carrying the avrB gene (Genetics 141:1597-1604). Plant tissue (expanded unifoliate leaves) was collected at 2, 4, 8, 12, 24, 36, and 53 hrs after inoculation and their mRNA pooled equally for cDNA construction. The library was prepared using the Stratagene Bluescript II SK(+) library construction kit. Complementary DNA was synthesized from mRNA using a primer consisting of a poly(dT) sequence with an XhoI restriction site. EcoRI adapters were ligated to the blunt-ended cDNA fragments followed by XhoI digestion. The cDNA insert is

protected from XhoI digestion via methylation during first strand synthesis. The cDNA fragments were directionally cloned into the EcoRI-XhoI restriction site of the Bluescript vector. The ligated cDNA fragments were transformed into E.coli Electromax DH10B host cells. Plant care, inoculations, and library construction were performed by Steve Clough (Ulla Vodka lab, University of Illinois).

BASE COUNT 147 a 108 c 135 g 179 t  
ORIGIN

## Alignment Scores:

Pred. No.:	4.7e-75	Length:	569
Score:	618.00	Matches:	123
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	100.00%	Indels:	0
DB:	12	Gaps:	0

US-09-786-715-8 (1-123) x BM177477 (1-569)

OY	1	MetAlaGluValGluGluGlyGluValIleGlyValHisThrValAspGluTrpLysLeu	20
DB	55	ATGGCTGAAGTGGAAAGAGGACAGGTCTATCGGCGTCCACACCGTTGATGAGTGAAGCTG	114
OY	21	GlnLeuGlnAsnAlaLysAspSerLysLeuIleValValAspPheThrAlaSerTrp	40
DB	115	CAACTCCAGATGCAAAAGACTCCAAAAAAGTGGTGGATTTACTGCTTCCCTGG	174
OY	41	CysGlyProCysArgPheMetAlaProValLeuAlaGluIleAlaLysLysThrProGlu	60
DB	175	TGTGGTCCATCCCGTTTATGCGCCCACTTCTTCGAGATTGCCAAGAAAAACCTCTGAA	234
OY	61	LeuIlePheLeuLysValAspValAspGluValArgProValAlaGluGluTyrSerIle	80
DB	235	TTGATCTTCTCCAAAGTGGATGTGATGAAGGCTGTGCTGAGGAATATTCAT	294
OY	81	GluAlaMetProThrPheLeuPheLeuLysAspGlyGluIleValAspLysValValGly	100
DB	295	GAGGCCATGCCAACCTTCTCTTGTGAAGATGGCGAGATCGTGACAGGTGGTGGT	354
OY	101	AlaSerLysAspAspLeuGlnAlaThrIleAlaLysHisAlaSerAlaValAlaAlaAla	120
DB	355	GCTAGTAGAGATGACCTTCAAGCCACCATAGCCAAAGCATGCTGTGTGCTGTCT	414
OY	121	SerSerSer 123	
DB	415	TCTTCTTCT 423	

Search completed: August 17, 2003, 23:34:27  
Job time : 1669.43 secs





TITLE cDNA expressed in Ricinus cotyledons  
JOURNAL Unpublished  
REFERENCE 2 (bases 1 to 603)  
AUTHORS Doljener E.  
TITLE Direct Submission  
JOURNAL Submitted (04-Apr-1996) Doljener E., University of Bayreuth,  
Department of Plant Physiology, Universitaetsstr.30, Bayreuth,  
Germany, 95440

FEATURES  
source location/Qualifiers  
1..603  
/organism="Ricinus communis"  
/mol\_type="mRNA"  
/cultivar="Sanguineus"  
/db\_xref="taxon:3988"  
/clone="PEDRH018"  
/tissue\_type="cotyledon"  
/clone\_lib="lambda Excel"  
/dev\_stage="seedling"  
23..379  
/codon\_start=1  
/product="thioredoxin"  
/protein\_id="CAA94534.1"  
/db\_xref="GI:1259594"  
/db\_xref="SPTREMBL:Q43636"  
/translation="MAEEGVIGCHTVEANNEOLQKNDKGLIVDFETASWCGPCR  
FIAPFLAEIAKLKLPVETFLKVDDELKTVAHEMAVESPTFELKGIIMDKVGAKK  
DELOOTIKHMAVAST"

CDS  
BASE COUNT 179 a 101 c 157 g 166 t  
ORIGIN

Alignment Scores:  
Pred. No.: 5.01e-44 Length: 603  
Score: 442.00 Matches: 80  
Percent Similarity: 87.61% Conservative: 19  
Best Local Similarity: 70.80% Mismatches: 14  
Query Match: 72.82% Indels: 0  
DB: 8 Gaps: 0

US-09-786-715-4 (1-118) x RCHTHIORXN (1-603)

OY 3 SerSerGIuGIuGLyGlnValIleGIcYshISserValAspGLuTrpLysGIuInphe 22  
Db 26 GCAGCAGAAAGAGGCAAGTGCATCGGCGCCACACTGTGAGCATGCAATGCAATG 85  
OY 23 GlnLysGIuValAspSerLysLysLeuValIleAspPheThAlaSerTrpLysGIu 42  
Db 86 CAGAAAGGAATGATACCAAGGACTGATCGTTGTGATTTACTGCTTCATGCTGTGCA 145  
OY 43 ProCysArgPheIleAlaProIleLeuAlaIuMetAlaLysLysThrProHisValIle 62  
Db 146 CCATGCCGTTTCATTCGTCCTTCCTGGCTGAGCGCCAGCAAACTGCCAAATGTACC 205  
OY 63 PheLeuLysValAspValAspGIuLeuLysThrValAlaGIuLupPheLysValIuAla 82  
Db 206 TTCCGAGAGGTGAGTGCATGACGTAAGACTGTTCCTACAGACTGGGCTGTGAGTCA 265  
OY 83 MetProThrPheValPheLeuLysGIuLysGIuValAlaGluArgLeuValAlaArg 102  
Db 266 ATGCCAACCTTATGTCCTCGAAGGAGGAGATTTATGACAAAGGTGGTGGCAAG 325  
OY 103 LysGIuLysLeuGlnAlaThrValGIuLysHisGIuAla 115  
Db 326 AAAGCAAGACTGCAGCAACTATACGCAACATGCGCT 364

RESULT 14  
AF323593 664 bp mRNA linear PLN 02-NOV-2001  
LOCUS AF323593  
DEFINITION Prunus persica thiodoxin H mRNA, complete cds.  
ACCESSION AF323593  
VERSION AF323593.1 GI:16588842  
KEYWORDS  
SOURCE Prunus persica (peach)  
ORGANISM Prunus persica

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
rosids; eurosids I; Rosales; Rosaceae; Amygdaloideae; Prunus.  
1 (bases 1 to 664)  
Callahan,A.M., Morgens,P.H. and Cohen,R.A.  
Isolation and initial characterization of cDNAs for mRNAs regulated  
during peach fruit development  
J. Am. Soc. Hortic. Sci. 118, 531-537 (1993)

REFERENCE  
JOURNAL 2 (bases 1 to 664)  
AUTHORS Callahan,A.M., Morgens,P.H., Cohen,R.A. and Scorza,R.  
TITLE Regulation of peach gene expression in a peach/almond hybrid  
JOURNAL Unpublished  
REFERENCE 3 (bases 1 to 664)  
AUTHORS Callahan,A.M., Dunn,L.L. and Cohen,R.A.  
TITLE Direct Submission  
JOURNAL Submitted (22-NOV-2000) USDA-ARS, Appalachian Fruit Research  
Station, 45 Wiltshire Rd., Kearneysville, WV 25430, USA

FEATURES  
source location/Qualifiers  
1..664  
/organism="Prunus persica"  
/mol\_type="mRNA"  
/cultivar="B612615"  
/db\_xref="taxon:3760"  
/clone="pch306c"  
/tissue\_type="ripe fruit"  
71..481  
/codon\_start=1  
/product="thioredoxin H"  
/protein\_id="AAI26915.1"  
/db\_xref="GI:16588843"  
/translation="MAEENOVIGCHTQAMEBOLKHNKLVVDFETASWCGCRL  
IAPILAEIAKLKPEVETFLKVDDELKTVAHEMAVESPTFELKGIIMDKVGAKKD  
ELQIVAKHVHVAAMAAASATSAATATATATASA"

CDS  
BASE COUNT 182 a 147 c 161 g 174 t  
ORIGIN

Alignment Scores:  
Pred. No.: 5.19e-43 Length: 664  
Score: 434.00 Matches: 81  
Percent Similarity: 84.35% Conservative: 16  
Best Local Similarity: 70.43% Mismatches: 18  
Query Match: 71.50% Indels: 0  
DB: 8 Gaps: 0

US-09-786-715-4 (1-118) x AF323593 (1-664)

OY 4 SerGIuGIuGLyGlnValIleGIcYshISserValAspGLuTrpLysGIuInphe 23  
Db 74 GCGAGGAAATCAAGTGCATCGGCGCCACACTCAAGCGGGAAGCAGCTCCAT 133  
OY 24 LysGIuValAspSerLysLysLeuValIleAspPheThAlaSerTrpLysGIu 43  
Db 134 AAGGAAAGCAAGAACAAAGTGGTGGGATTTCAAGCTTCGCTGTGAGACCG 193  
OY 44 CysArgPheIleAlaProIleLeuAlaIuMetAlaLysLysThrProHisValIlePhe 63  
Db 134 TCCCGGTGATCGCCCATCTTCGCGGAGTGTGCTTAAGAACCCAGAAAGTACGTTG 253  
OY 64 LeuLysValAspValAspGIuLeuLysThrValAlaGIuLupPheLysValIuAlaMet 83  
Db 254 CTAAAGGTGAGAGTGCATGACGTAAGACTGTTCCTCGAGAGTGGGCTGTGAGCAATG 313  
OY 84 ProThrPheValPheLeuLysGIuLysGIuValAlaGluArgLeuValAlaArg 103  
Db 314 CCTACCTTCCTTCCTCAAGCAAGCAAGATGAGCAAGGTTGTGGTGCAGAA 373  
OY 104 GluLysLeuGlnAlaThrValGIuLysHisGIuAla 118  
Db 374 GACGAGTACAGATCAAAAGTTGCCAAGCATGTGCGCGCCGCC 418

RESULT 15  
PSA319808 593 bp mRNA linear PLN 17-JUN-2001  
LOCUS PSA319808

NTTNA  
LOCUS NTTNA 698 bp mRNA linear PLN 15-FEB-1994  
DEFINITION N. tabacum mRNA for thioredoxin.  
ACCESSION X58527  
VERSION X58527.1 GI:20046  
KEYWORDS thioredoxin.  
SOURCE Nicotiana tabacum (common tobacco)  
ORGANISM Nicotiana tabacum  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
Asteridae; lamids; Solanales; Solanaceae; Nicotiana.  
REFERENCE  
AUTHORS Marty, I. and Meyer, Y.  
TITLE Nucleotide sequence of a cDNA encoding a tobacco thioredoxin  
JOURNAL Plant Mol. Biol. 17 (1), 143-147 (1991)  
MEDLINE 91329721  
PubMed 1868216  
REFERENCE  
AUTHORS 2 (bases 1 to 698)  
TITLE Brugidou, C., Marty, I., Chartier, Y. and Meyer, Y.  
JOURNAL The Nicotiana tabacum genome encodes two cytoplasmic thioredoxin  
MEDLINE 93241165  
PubMed 8479434  
REFERENCE  
AUTHORS 3 (bases 1 to 698)  
TITLE Meyer, Y.  
JOURNAL Direct Submission  
MEDLINE  
PubMed  
REFERENCE  
AUTHORS Submitted (14-FEB-1991) Y. Meyer, Laboratoire de Physiologie et  
JOURNAL Biologie, Moleculaire Vegetale, Universite av de Villeneuve, 66860  
COMMENT Gene product is probably cytoplasmic.  
FEATURES  
source location/Qualifiers  
1. 698  
/organism="Nicotiana tabacum"  
/mol\_type="mRNA"  
/strain="white Burley"  
/db\_xref="taxon:4097"  
/clone\_lib="plasmid pTZ19"  
/dev\_stage="in vitro cells"  
1. 682  
/evidence="experimental"  
79.459  
/codon\_start=1  
/product="thioredoxin"  
/protein\_id="CAA41415.1"  
/db\_xref="GI:20047"  
/db\_xref="SWISS-PROT:P29449"  
/translation="MAANDATSSSEGCQVFGCHKVEEMNEFFKGVETKKLVVDFTAS  
WCGPCRFIADIAIKMPHVFILKVDDELKTVSAEWSVEAMPTEFVTKDKEVDR  
VWGAKKEELQOTIVKHAAPATVTA"  
polya\_signal 468.473  
polya\_signal 511.517  
BASE COUNT 192 a 117 c 164 g 225 t  
ORIGIN

Alignment Scores:  
Pred. No.: 3,44e-49 length: 698  
Score: 485.50 Matches: 90  
Percent Similarity: 89.08% Conservative: 16  
Best Local Similarity: 75.63% Mismatches: 10  
Query Match: 79.98% Indels: 3  
DB: Gaps: 1

US-09-786-715-4 (1-118) x NTTNA (1-698)

QY 3 SerSerGluGluGlyGlnValIleGlyCysHisSerValAspGluTrpLysGluGlnPhe 22  
|||||  
Db 100 TCATCCGAGAGGAGGACAGTGTCCGCTGCCACAGGTGAGGAATGGAACGAGTACTTC 159  
GlnLysGlyValAspSerLysLysLeuValValIleAspPheThrAlaSerTrpCysGly 42  
|||||  
Db 160 AAGAAAGGCGTGTGACTAAGAAAGTGTGTGTGCTGATTTTACTCTCTCATGTCGCGC 219

QY 43 ProCysArpPheIleAlaProIleLeuAlaGluMetAlaLysTrpProHisValIle 62  
|||||  
Db 220 CCTTGGCGTTTATFTGCCCAATCTTCTGCGACATTGCTAAGAAAGTGGCCCATGTATA 279  
PheLeuLysValAspValAspGluLeuLysTrpValAlaGluGluPheLysValGluAla 82  
|||||  
Db 280 TTCCTCAGGTTGATGTGATGATGAACGACGCTTTCACCGCAATGAGTGTGGAGGCA 339  
MetProThrPheValPheLeuLysGluGlyLysGluValGluLysGluValGluAlaArg 102  
|||||  
Db 340 ATGCCACATTCTTCTCTCATTAAGATGGAAGAAAGAGTGACAGAGTGTGGTGCACA 399  
LysGluGluLeuGlnAlaThrValGluLysHisGlyAla-----IleThrAla 118  
|||||  
Db 400 AAGAGAGAGTTGACAGACAGACCATGAGACATGTCCTCTGCTACTGCTACTGCT 456

RESULT 2  
LOCUS AR016869 653 bp DNA linear PAT 05-DEC-1998  
DEFINITION Sequence 103 from patent US 5777200.  
ACCESSION AR016869  
VERSION AR016869.1 GI:3973146  
KEYWORDS  
SOURCE Unknown.  
ORGANISM Unknown.  
REFERENCE  
AUTHORS 1 (bases 1 to 653)  
TITLE Ryals, J.A., Alexander, D.C., Goodman, R.M. and Stinson, J.R.  
Chemically regulatable and anti-pathogenic DNA sequences and uses thereof  
JOURNAL Patent: US 5777200-A 103 07-JUL-1998;  
FEATURES  
source location/Qualifiers  
1. 653  
/organism="unknown"  
BASE COUNT 165 a 114 c 157 g 215 t 2 others  
ORIGIN

Alignment Scores:  
Pred. No.: 8.84e-48 length: 653  
Score: 473.50 Matches: 90  
Percent Similarity: 88.33% Conservative: 16  
Best Local Similarity: 75.00% Mismatches: 10  
Query Match: 78.01% Indels: 4  
DB: Gaps: 1

US-09-786-715-4 (1-118) x AR016869 (1-653)

QY 3 SerSerGluGluGlyGlnValIleGlyCysHisSerValAspGluTrpLysGluGlnPhe 22  
|||||  
Db 69 TCATCCGAGAGGAGGACAGTGTCCGCTGCCACAGGTGAGGAATGGAACGAGTACTTC 128  
GlnLysGlyValAspSerLysLysLeuValValIleAspPheThrAlaSerTrpCysGly 42  
|||||  
Db 129 AAGAAAGGCGTGTGACTAAGAAAGTGTGTGTGCTGATTTTACTCTCTCATGTCGCGC 188  
ProCysArpPheIleAlaProIleLeuAlaGluMetAlaLysTrpProHisValIle 62  
|||||  
Db 189 CCTTGGCGTTTATFTGCCCAATCTTCTGCGACATTGCTAAGAAAGTGGCCCATGTATA 248  
PheLeuLysValAspValAspGluLeuLysTrpValAlaGluGluPheLysValGluAla 82  
|||||  
Db 249 TTCCTCAGGTTGATGTGATGATGAACGACGCTTTCACCGCAATGAGTGTGGAGGCA 308  
MetProThrPheValPheLeuLysGluGlyLysGluValGluLysGluValGluAlaArg 102  
|||||  
Db 309 AATGCCACATTCTTCTCTCATTAAGATGGAAGAAAGAGTGACAGAGTGTGGTGCACA 368  
GlnLysGluGluLeuGlnAlaThrValGluLysHisGlyAla-----IleThrAla 118  
|||||  
Db 369 GAAAGAGAGTTGACAGACAGACCATGAGACATGTCCTCTGCTACTGCTACTGCT 426

RESULT 3

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - nucleic search, using frame\_plus\_p2n model

Run on: August 17, 2003, 18:47:50 ; Search time 1825.46 Seconds

(without alignments)  
2644.453 Million cell updates/sec

Title: US-09-786-715-4

Perfect score: 607

Sequence: 1 MASSEGGVICHSHVDENK.....YGARRELQATVEKGAITA 118

Scoring table:

BLOSUM62	
Xgapop 10.0 , Xgapext 0.5	
Ygapop 10.0 , Ygapext 0.5	
Fgapop 6.0 , Fgapext 7.0	
Delop 6.0 , Delext 7.0	

Searched: 2888711 seqs, 20454813386 residues

Total number of hits satisfying chosen parameters: 5777422

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+ p2n model -DEV=xlp  
-Q=/ggn2.1/USPTO/us09786715/runat\_11082003\_150514\_6035/app-query.fasta\_1.1052  
-DB=GenEmbl -QFMT=fastcap -SUFFIX=rge -MINMATCH=0.1 -LOOPEL=0 -LOOPEXT=0  
-UNITS=bits -START=1 -END=1 -MATRIX=biosum62 -TRANS=human4.0.cdi -LIST=45  
-DOCALIGN=200 -THR.SCORE=pct -THR.MAX=100 -THR.MIN=0 -ALIGN=15 -MODE=LOCAL  
-OUTFMT=plc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=200000000  
-USER=US09786715.ecgn.1.1.5721.0runat\_11082003\_150514\_6035 -NCPU=6 -ICPU=3  
-NO\_MMAP -LARGESQUERY -NEG.SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG  
-DEV.TIMEOUT=120 -WARN.TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6  
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

GenEmbl:\*  
1: gb\_ba:\*  
2: gb\_hcg:\*  
3: gb\_in:\*  
4: gb\_om:\*  
5: gb\_ov:\*  
6: gb\_pat:\*  
7: gb\_ph:\*  
8: gb\_pl:\*  
9: gb\_pr:\*  
10: gb\_ro:\*  
11: gb\_sts:\*  
12: gb\_sy:\*  
13: gb\_un:\*  
14: gb\_vt:\*  
15: em\_da:\*  
16: em\_fun:\*  
17: em\_hum:\*  
18: em\_in:\*  
19: em\_mu:\*  
20: em\_om:\*  
21: em\_or:\*  
22: em\_ov:\*  
23: em\_pat:\*  
24: em\_ph:\*  
25: em\_pl:\*  
26: em\_ro:\*  
27: em\_sts:\*  
28: em\_un:\*

29: em\_vt:\*  
30: em\_hcg\_hum:\*  
31: em\_hcg\_inv:\*  
32: em\_hcg\_other:\*  
33: em\_hcg\_mus:\*  
34: em\_hcg\_pln:\*  
35: em\_hcg\_rtd:\*  
36: em\_hcg\_mam:\*  
37: em\_hcg\_vrt:\*  
38: em\_sy:\*  
39: em\_hcgo\_hum:\*  
40: em\_hcgo\_mus:\*  
41: em\_hcgo\_other:\*

Prod. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	485.5	80.0	698	8	NITRNA	X58527 N.tabaccu m
2	473.5	78.0	653	6	AR016869	AR016869 Sequence
3	473.5	78.0	653	6	AR020895	AR020895 Sequence
4	473.5	78.0	653	6	AR027218	AR027218 Sequence
5	473.5	78.0	653	6	AR038505	AR038505 Sequence
6	473.5	78.0	653	6	AR064647	AR064647 Sequence
7	473.5	78.0	653	6	AR067572	AR067572 Sequence
8	473.5	78.0	653	6	I38524	I38524 Sequence 10
9	473.5	78.0	653	6	I56999	I56999 Sequence 10
10	473.5	78.0	653	6	I59865	I59865 Sequence 10
11	473.5	78.0	653	6	I75192	I75192 Sequence 10
12	444	73.1	665	8	AY170650	AY170650 Plasmid set
13	442	72.8	603	8	RCYH10RXN	Z70677 R.communis
14	434	71.5	664	8	AF323593	AF323593 Prunus pe
15	432	71.2	593	8	PSA319808	AJ319808 Plasmid sat
16	425	70.0	595	8	D87984	D87984 Fagopyrum e
17	422	69.5	630	8	PSA310990	AJ310990 Plasmid sat
18	419	69.0	345	6	AX505468	AX505468 Sequence
19	419	69.0	480	8	AY088687	AY088687 Arabidops
20	419	69.0	497	8	ATPH10ARA	Z14084 A.thaliana
21	419	69.0	784	8	AY271308	AY271308 Citrus x
22	406	66.9	345	8	AF483265	AF483265 Populus t
23	392	64.6	357	8	AY040028	AY040028 Arabidops
24	392	64.6	528	8	ATTH1RED4	Z35476 A.thaliana
25	392	64.6	530	8	AK118035	AK118035 Arabidops
26	392	64.6	560	8	AY087159	AY087159 Arabidops
27	392	64.6	561	8	AF360227	AF360227 Arabidops
28	390	64.3	526	11	G73679	G73679 R2488R elio
29	390	64.3	601	6	OSU92541	U92541 Oryza sativ
30	390	64.3	686	6	E08194	E08194 RPS13 gene
31	390	64.3	687	8	RIC1H	D21836 Oryza sativ
32	389	64.1	357	6	AX412276	AX412276 Sequence
33	389	64.1	357	6	AX412277	AX412277 Sequence
34	389	64.1	357	6	AX412278	AX412278 Sequence
35	389	64.1	357	6	AX412279	AX412279 Sequence
36	389	64.1	357	6	AX412473	AX412473 Sequence
37	389	64.1	357	6	AX412474	AX412474 Sequence
38	389	64.1	357	6	AX412566	AX412566 Sequence
39	389	64.1	357	6	AX412917	AX412917 Sequence
40	389	64.1	357	6	AX507041	AX507041 Sequence
41	389	64.1	357	6	AX651357	AX651357 Sequence
42	389	64.1	402	8	AY114566	AY114566 Arabidops
43	389	64.1	440	8	AY093318	AY093318 Arabidops
44	389	64.1	469	8	AY065098	AY065098 Arabidops
45	389	64.1	489	8	ATTH1RED2	Z35474 A.thaliana

RESULT 1

#### ALIGNMENTS

recombinant cDNAs were transformed into E. coli XL1-Blue cells (Clontech). Suspensor cDNA plasmids used for directional sequencing were obtained by in vivo excision from the lambda Triplex2 recombinants in E. coli BM25.8 cells (Clontech).

BASE COUNT 154 a 92 c 134 g 190 t  
ORIGIN

Alignment Scores:

Pred. No.:	8,31e-61	Length:	570
Score:	533.00	Matches:	100
Percent Similarity:	95.00%	Conservative:	14
Best Local Similarity:	83.33%	Mismatches:	6
Query Match:	86.67%	Indels:	0
DB:	14	Gaps:	0

US-09-786-715-6 (1-120) x CA901886 (1-570)

QY	1	MetaIaglySerSerGIuGIuValIleSerCysHisThrValGIuGIuTrpAsn	20
DB	82	ATGGCCGGCGCATCAGAGGAGGAGCAAGTCATTCCTGCCACACCGTTGAGCGCATGACA	141
QY	21	AspGlnLeuGlnIlySgIyAsnGluSerIlySleuIleValIaAspPheThrAlaSer	40
DB	142	GACCACTCGAAAGGCGCAATTAATCCAGAACTAATGTTGTGATTTACTGCTTCT	201
QY	41	TrpCysGlyProCysArgPheIleAlaProPheLeuAlaGIuLeuAlaIlySlyAsPheThr	60
DB	202	TGGTGTGACCATGCCGTTTCATTTCTTCATTTGGCTGAGCTGGCTAGAGAGTTTACA	261
QY	61	SerValIlePheLeuIlySValAspValAspGluLeuIlySerValSerGlnAspTrpAla	80
DB	262	AATGTGTAATTCGAGAGGTGATGATGACGAATTAAGAGTGTGCTGAAGATTTTGTCT	321
QY	81	IleGluAlaMetProThrPheValPheValIlySgIuGIuThrLeuLeuAspIlyValIal	100
DB	322	GTTGAGGCAATGCCAATCTTTGTGTGTGTAAGAGGGAACCTCTTGGGCAAGTGCTG	381
QY	101	GlyAlaIlySlyAspGluLeuGlnIlySleuIleGlnIlySHisValAlaSerAlaSerAla	120
DB	382	GGACCAAGAAAGAAAGATTGACAGACACATAGAGAAACATGTGCTGCACCTAGTGTCT	441

Search completed: August 17, 2003, 23:34:21  
Job time : 1628.86 secs

OY 114 sva1a1aserAlaserA1a 120  
 ||||||||||||||||  
 Db 201 TGTGCTTCAGCTAGTCT 183

RESULT 14  
 CD038084 687 bp mRNA linear EST 07-MAY-2003  
 LOCUS UPP1002.C04 USDA-Tifton Peanut Immature Pod cDNA Library (UTPP)  
 DEFINITION Arachis hypogaea cDNA clone UTPP1002.C04 5', mRNA sequence.  
 ACCESSION CD038084  
 VERSION CD038084.1 GI:30419922  
 SOURCE EST.  
 ORGANISM Arachis hypogaea (peanut)  
 Arachis hypogaea  
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids  
 ; eurosoids I; Fabales; Fabaceae; Papilionoideae; Aeschynomeneae;  
 Arachis.

REFERENCE 1 (bases 1 to 687)  
 AUTHORS Luo, M., Dang, P., Guo, B. Z., Holbrook, C. C., Lee, R. D., Bauscher, M. G.  
 TITLE Generation and Analyses of ESTs for Arachis hypogaea  
 JOURNAL Unpublished  
 COMMENT Contact: Baozhu Guo  
 Molecular Genetics  
 USDA/ARS, Crop Protection and Management Research Unit  
 2747 Davis Rd., Tifton, GA 31794, USA  
 Tel: 229-387-2334  
 Fax: 229-387-2321  
 Email: bguo@tifton.usda.gov  
 Seq primer: T3.

FEATURES  
 source location/Qualifiers  
 1..687  
 /organism="Arachis hypogaea"  
 /mol\_type="mRNA"  
 /cultivar="A13"  
 /db\_xref="taxon:3818"  
 /clone="UTPP1002.C04"  
 /tissue\_type="Immature pods"  
 /dev\_stage="R6"  
 /lab\_host="X11-blue"  
 /clone\_lib="USDA-Tifton Peanut Immature Pod cDNA Library (UTPP)"  
 /note="Vector: Uni-ZAP XR; Site.1: EcoRI; Site.2: XhoI; cDNA library was constructed from peanut cultivar A13 (NCV1XAR4). A13 has resistance to Aspergillus infection and drought tolerance. The immature pods that developed to R6 stage were collected from different plants, and placed into liquid N2 immediately and stored in -80oc freezer. Total RNA was isolated with TRIzol-Reagent ultrapure(GIBCOBRL). mRNA was extracted and purified from total RNA (Promega). cDNA synthesis and library construction followed the protocol of by ZAP-cDNA GigaPack II Gold Cloning Kit (Stratagene). The cDNA above 500bp were collected after size-fraction. The inserts were directionally cloned into Uni-ZAP XR vector using XhoI EcoRI sites adapters. The lambda library was packed into phages using GigaPack III Gold (Stratagene). The un-amplified library was used to excise plunescrpt phagemids from the Uni-ZAP XR vector, and the phagemids was used to transform the host bacteria SOLR. The library was constructed by Dr. Meng Luo and Dr. Phat Dang."

BASE COUNT 189 a 107 c 176 g 215 t  
 ORIGIN

Alignment Scores:  
 Pred. No.: 8..92e-64 Length: 687  
 Score: 556.00 Matches: 103  
 Percent Similarity: 96.67% Conservative: 13  
 Best Local Similarity: 85.83% Mismatches: 4  
 Query Match: 90.41% Indels: 0  
 DB: 14 Gaps: 0

US-09-786-715-6 (1-120) x CD038084 (1-687)

OY 1 MetAlaGlySerSerGluGluGluGlnValIleSerCysHisThrValGluGluTyrPsn 20  
 |||||||  
 Db 69 ATGGCTGCTTCATCAGAGAGGAGGACAGTTATGCTTCCCTGCTGATGATGACACC 128

OY 21 AspG1nLeuGlnGlnGlnGlnGlnSerValSerValSerValSerValSerValSer 40  
 |||||||  
 Db 129 CAACAGCTCGAGAGAGGAATGATCTAAGACCTGATTTGTGATTTGCTGCTTCA 188

OY 41 TrpCysG1nProCysArpPheIleAlaProPheLeuAlaGluLeuAlaLysLysPheThr 60  
 |||||||  
 Db 189 TGTGCGGACCATGCGCTTCATTCGACCTTTCTGCTGAGTGGCTAAGAGATTACA 248

OY 61 SerValIlePheLeuLysValAspValAspGluLeuLysSerValSerGlnAspTyrPala 80  
 |||||||  
 Db 249 AATGTCATATTTTGAAGGTGAGTGTGAGTGAACCTGACGCTCTCTCAAGCTGGCT 308

OY 81 IleGluAlaMetProThrPheValPheValIleGluG1nThrLeuLeuAspLysValVal 100  
 |||||||  
 Db 309 GTGGAGGACATGGCCACCTTGTGTTGTGTAAGAGGAAACATTTATGCAAGTGC 368

OY 101 GlyAlaLysLysAspGluLeuGlnGlnLysIleGlnLysHisValAlaSerAla 120  
 |||||||  
 Db 369 GGAGCAAG 428

RESULT 15  
 CA901886 570 bp mRNA linear EST 27-DEC-2002  
 LOCUS PCSC14170 Scarlet Runner Bean Suspensor Region Triplex2 Phaseolus  
 DEFINITION coccineus cDNA 5' similar to Thoredoxin H-type, mRNA sequence.  
 ACCESSION CA901886  
 VERSION CA901886.1 GI:27388878  
 KEYWORDS EST.  
 ORGANISM Phaseolus coccineus  
 Phaseolus coccineus  
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids  
 ; eurosoids I; Fabales; Fabaceae; Papilionoideae; Phaseoleae;  
 Phaseolus.

REFERENCE 1 (bases 1 to 570)  
 AUTHORS Bul, A.O., Le, B. H., Weterings, K., Bi, Y., -P., Choi, J. -S., McElroy, K. E.  
 TITLE Gene Activity in Different Regions of a Post-Fertilization Plant  
 JOURNAL Embryo by EST Analysis  
 COMMENT Unpublished  
 CONTACT: Goldberg, R. B.  
 Department of Molecular, Cell, & Developmental Biology  
 University of California, Los Angeles  
 621 Charles E. Young Drive South, Los Angeles, CA 90095-1606, USA  
 Tel: 310 825 3270  
 Fax: 310 825 8201  
 Email: bob@ucla.edu  
 Seq primer: 5' Triplex  
 POLYA-NO.

FEATURES  
 source location/Qualifiers  
 1..570  
 /organism="Phaseolus coccineus"  
 /mol\_type="mRNA"  
 /cultivar="Hammond's Dwarf Scarlet"  
 /db\_xref="taxon:3886"  
 /dev\_stage="6-days post-pollination"  
 /clone\_lib="Scarlet Runner Bean Suspensor Region Triplex2"  
 /note="Organ: Suspensor Region of Globular Stage Embryos; Vector: Triplex2; Site.1: SfiI; Site.2: SfiI; Suspensor regions were micro-dissected from globular-stage embryos six days after pollination from greenhouse-grown plants [Weterings et al., Plant Cell 13, 2409-2425 (2001)]. Double-stranded cDNA was synthesized from suspensor mRNA using the SMART cDNA Library Construction Kit according to the manufacturer (Clontech). The suspensor cDNA fragments were directionally ligated into the SfiI restriction site of the lambda Triplex2 vector (Clontech), and the



```

DB      355 GGAGCAAGAGATGAGCTGCAGCAGACAAATATGCTGCTCAGCTAGTGCCT 414
|||||
RESULT 11
LOCUS   AM569018
DEFINITION
  466 bp mRNA linear EST 03-DEC-2001
  s174e02.y1 Gm-cl031 Glycine max cDNA clone GENOME SYSTEMS CLONE ID:
  Gm-cl031-435 5' similar to TR:Q43636 Q43636 THIOREDOXIN. ;, mRNA
  sequence.
ACCESSION
  AM569018
VERSION
  AM569018.1 GI:7233673
KEYWORDS
  EST.
SOURCE
  Glycine max (soybean)
  Glycine max
  Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
  Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids
  ; eurosids I; Fabales; Fabaceae; Papilionoideae; Phaseoleae;
  Glycine.
REFERENCE
  1 (bases 1 to 466)
  Shoemaker,R., Keim,P., Vodkin,L., Erpelting,J., Corryell,V., Khanna
  ,A., Bolla,B., Matra,M., Hillier,L., Kucaba,T., Martin,J., Beck,C.,
  Wylie,T., Underwood,K., Steptoe,M., Theising,B., Allen,M., Bowers
  ,Y., Person,B., Swaller,T., Gibbons,M., Pape,D., Harvey,N., Schurk
  ,R., Rittner,E., Kohn,S., Shin,T., Jackson,Y., Cardenas,M., McCann
  ,R., Materston,R. and Wilson,R.
  Public Soybean EST Project
  Unpublished
  Contact: Shoemaker R/Public Soybean EST Project
  Washington University School of Medicine
  4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA
  Tel: 314 286 1800
  Fax: 314 286 1810
  Email: est@watson.wustl.edu
  This clone is available through: Resgen, Invitrogen Corp. 2130
  South Memorial Parkway Huntville, AL 35801 For further information
  call: (800)-533-4363 or contact via email: cou@resgen.com
  High quality sequence stop: 416.
FEATURES
  source
  1..466
    /organism="Glycine max"
    /mol_type="mRNA"
    /db_xref="taxon:3847"
    /clone="GENOME SYSTEMS CLONE ID: Gm-cl031-435"
    /tissue_type="Williams seedlings, minus the cotyledons"
    /lab_host="DH10B"
    /clone_id="Gm-cl031"
    /note="Vector: pBluescript II XR; Site_1: EcoRI; Site_2:
    XhoI; This cDNA library was constructed from mRNA isolated
    from whole 'Williams' seedlings, minus the cotyledons,
    which were propagated on paper towels with distilled water
    for 5 days, incubated at 40 degrees C for 1 hour. The
    cotyledons were removed and the remaining tissue was
    flash-frozen in liquid nitrogen. Stratagene's cDNA
    synthesis kit (catalog number 200401) was used to
    synthesize the cDNA. First-strand synthesis was performed
    with 5-methyl dcp, hence the ligated cDNA was
    hemimethylated. A modification of Stratagene's
    first-strand synthesis primer was used. An anchor'
    nucleotide (V-A, C, or G) was added to the 3' end of the
    primer [GAGAGAGAGAGAGAGAGAGACTGTCGAG(T)18V] to anchor
    the primer at the 5' end of the poly(A) tract. After
    second strand synthesis, the cDNA ends were filled in with
    cloned pfu DNA, ligated to EcoRI adapters and
    subsequently phosphorylated. The cDNA was then
    precipitated and redissolved in sterile, RNase-,
    DNase-free water. The XhoI site within the first-strand
    synthesis primer was then restricted by digestion with
    XhoI from Promega (400/ul); all XhoI sites in the cDNA
    would be protected by their hemimethylated status. The
    cDNA constructs were size-fractionated with a 500 bp
    cutoff, using GIBCOBRL Life Technologies' cDNA size
    fractionation column. The column eluent was then
  
```

```

BASE COUNT 142 a 82 c 116 g 126 t
ORIGIN
Alignment Scores:
Pred. No.: 2,39e-67 Length: 466
Score: 581.00 Matches: 112
Percent Similarity: 97.508 Conservative: 5.
Best Local Similarity: 93.33% Mismatches: 3
Query Match: 94.47% Indels: 0
Gaps: 0
DB:
US-09-786-715-6 (1-120) x AM569018 (1-466)
QY 1 MetAlaGlySerSerGluGluGluGluValIleSerCysHisThrValGluUtrPasn 20
  |||||
  69 ATGGCAGGCTCCTCGAGAGAGGAGACAAATGATTAGCTGCACACCGTTGACCATGGAC 128
  |||||
QY 21 AspGlnLeuGlnIlySerGlnSerGlnSerGlnSerGlnSerGlnSerGlnSer 40
  |||||
  129 GATCAACTCCAGAAAGGCAACCAATCCAAATCTATGTTGATTTTCTGCTTCT 188
  |||||
QY 41 TrpCysGlyProCysArgPheIleAlaProPheLeuAlaGlnLeuAlaLysLysPheThr 60
  |||||
  189 TGGTGTGAGCAATGCCGTTTTCATTCGACCATTTCTTGCTGACCTGCTAAGAGTTCA 248
  |||||
QY 61 SerValIlePheIleuLysValLysPheValLysPheValLysPheValLysPheVal 80
  |||||
  249 AGTGTGCTATTCCTTAAGAGTGATGAGCAATTAAGAGTGTCTCAAGACTGGGCT 308
  |||||
QY 81 IleGluAlaMetProThrPheValIlePheValIleGluGluGluGluGluGluGlu 100
  |||||
  309 ATTTGAGGCTATGCCCACTTTTGTGTTGTAAGAGGAACTTTTAAGCAAGTGGG 368
  |||||
QY 101 GYAAlaLysLysAspGluLeuGlnIlyIleGlnIlyHisValAlaSerAlaSerAla 120
  |||||
  369 GGAGCAAGAGATGAGCTGCAGCAGACAAATATGCTGCTCAGCTAGTGCCT 428
  |||||
RESULT 12
LOCUS   BG838879/c
DEFINITION
  613 bp mRNA linear EST 25-MAY-2001
  Gc01_06f08_R Gc02_AAFc_ECORc_cold_stressed Glycine_clandestina_SSH
  Glycine clandestina cDNA clone Gc01_06f08, mRNA sequence.
ACCESSION
  BG838879
VERSION
  BG838879.1 GI:14205201
KEYWORDS
  EST.
  Glycine clandestina
  Glycine clandestina
  Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
  Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids
  ; eurosids I; Fabales; Fabaceae; Papilionoideae; Phaseoleae;
  Glycine.
REFERENCE
  1 (bases 1 to 613)
  Singh,J.A., Savole,A.J., Chapados,J., Courroux,P., De Moors,A.,
  Harris,L.J., Hattori,J.I., Ouellet,T., Robert,L.S., Spott,D. and
  Tinker,N.A.
  Differentially Expressed Sequence Tags from Cold-Stressed Glycine
  clandestina Seedlings
  Unpublished
  Contact: Singh,J.A.
  Eastern Cereal and Oilseed Research Centre
  Agriculture and Agri-food Canada
  KW Neatby Bldg., Central Experimental Farm, Ottawa, Ontario, K1A
  0C6, Canada
  Tel.: (613) 759-1662
  Fax: (613) 759-1701
  precipitated, redissolved, and ligated into Stratagene's
  pBluescript II XR predigested vector (pBluescript II SK(+))
  vector that has been digested with EcoRI and XhoI, and
  phosphorylated by Stratagene). 100% of the white and blue
  colonies appear to contain recombinant plasmids with cDNA
  inserts, based on size (7-18 and 5, respectively). This
  library was constructed by Dr. Paul Keim and Dr. Virginia
  Corryell."
  
```



material was provided by Michael G. Hahn (Complex Carbohydrate Research Center, University of Georgia) and the library was constructed by Anu Khanna (Lila Vodkin lab, University of Illinois).

BASE COUNT 142 a 97 c 120 g 163 t  
ORIGIN

## Alignment Scores:

Pred. No.: 2,05e-67 Length: 522  
Score: 582.00 Matches: 113  
Percent Similarity: 100.008 Conservative: 0  
Best Local Similarity: 100.008 Mismatches: 0  
Query Match: 94.63% Indels: 0  
DB: 13 Gaps: 0

US-09-786-715-6 (1-120) x BQ786610 (1-522)

QY 8 GLYGINVALLIISERCYSHSTHVRVALGLUTRPAASPLINLEUGINLYSGLYASN 27  
DB 1 GGACAGGTCAATTAGCTGCCACACCGTTGAAGATGACATCACTCCAGAGGCGAC 60  
QY 28 GLUSERTYSLYSLEULI VALASPHETHRALASERTPCYSGIYPCYCYARGPHE 47  
DB 61 GAATCCAGAACTCATTTGTTGTGATTTACTGCTTCTGTGTGGACCATGCCGTTTC 120  
QY 48 ILEALAPROPHLEUALAGLULEUALYLSYPHERHSERYALLIHPHEULYSVAL 67  
DB 121 ATTGCACCATTTGGCTGAGTGGCTAGAGAGTTCACAGTGCATATTCCTAAAGGTG 180  
QY 68 ASYVALASPLULEULYSSEVALSERGLASPTTRPALLEGUALMETPROTHRPH 87  
DB 181 GAGTGTGACAAATTAAGAGCTGTTCTCAAGATTGGCTATGAGGCTATGCCACTTTT 240  
QY 88 VALPHEVALLYSGIUGLYTHRIEULASPLYVALVALGLVALYLSYASPLULEU 107  
DB 241 GTGTTTGTGAAGAGGACGCTTCTGCAAAAGTGGTGGACCAAGAGATGACCTG 300  
QY 108 GLINGLYSLIEGLINLYSHISVALASERVALSERVAL 120  
DB 301 CACGAGAAATACAGAAACATGTGGCTTCAGTACTGCT 339

RESULT 10  
LOCUS A1988470 452 bp mRNA linear EST 30-NOV-2001  
DEFINITION sc02107.y1 Gm-c1020 Glycine max cDNA clone GENOME SYSTEMS CLONE ID: Gm-c1020-542 5' similar to TR:Q43636 Q43636 THIOREDOXIN. ; mRNA sequence.  
ACCESSION A1988470  
VERSION A1988470.1 GI:5820264  
KEYWORDS EST.  
SOURCE Glycine max (soybean)  
ORGANISM Glycine max

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids; eustoids I; Fabales; Fabaceae; Papilionoideae; Phaseoleae; Glycine.

REFERENCE 1 (bases 1 to 452)  
AUTHORS Shoemaker, R., Keim, P., Vodkin, L., Erpelting, J., Corvelli, V., Khanna, A., Bolla, B., Marra, M., Hillier, L., Kucaba, T., Martin, J., Beck, C., Wylie, T., Underwood, K., Stepien, M., Theising, B., Allen, M., Bowers, Y., Person, B., Swaller, T., Gibbons, M., Pape, D., Harvey, N., Schurk, R., Ritter, E., Kohn, S., Shin, T., Jackson, Y., Cardenas, M., McCann, R., Waterston, R. and Wilson, R.  
TITLE Public soybean EST Project  
JOURNAL Unpublished

COMMENT Contact: Shoemaker R/Public Soybean EST Project  
Public Soybean EST Project  
Washington University School of Medicine  
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA  
Tel: 314 286 1800  
Fax: 314 286 1810  
Email: est@watson.wustl.edu  
This clone is available through: Resgen, Invitrogen Corp. 2130

South Memorial Parkway Huntsville, AL 35801 For further information call: (800)-533-4363 or contact via email: cou@resgen.com  
Seq primer: -40RP from Gibco  
High quality sequence stop: 413.  
Location/Qualifiers

## FEATURES

## source

1.452  
/organism="Glycine max"  
/mol\_type="mRNA"  
/db\_xref="taxon:3847"  
/clone="GENOME SYSTEMS CLONE ID: Gm-c1020-542"  
/tissue\_type="root nodules of greenhouse grown plants"  
/lab\_host="X110-Gold"  
/clone\_lib="Gm-c1020"  
/note="Vector: pBluescript II SK+; Site 1: EcoRI; Site 2: XhoI; This cDNA library was constructed from mRNA isolated from nodules on the roots of 2.5 month-old Glycine max 'Williams' plants that were greenhouse grown. The cDNA library was prepared using the Stratagene pBluescript II SK(+) library construction kit. First strand synthesis was performed with 5-methyl dCTP, hence the ligated cDNA was hemimethylated. A modification of Stratagene's first-strand synthesis primer was used. An 'anchor' nucleotide (V-A, C, or G) was added to the 3' end of the primer [GAGACAGACAGACAGACAGACTGCTCGAG(T)18V] to anchor the primer at the 5' end of the poly(A) tract. After second-strand synthesis, the cDNA ends were filled in with cloned Pfu DNA polymerase and size-fractionated with a 400 bp cutoff, using a Siesep 400 Spin column from Pharmacia. The column eluent was ligated to EcoRI adaptors and phosphorylated. The XhoI sites in the cDNA would be protected by their hemimethylated status. The cDNA constructs were size-fractionated with a 500 bp cutoff, using GibcoBRL Life Technologies' cDNA size predestiged vector (pBluescript II SK+) that has been digested with EcoRI and XhoI, and phosphorylated by Stratagene. Both the white and blue colonies appear to contain recombinant plasmids with cDNA inserts, based on size (n=56) and sequence (n=16). This library was constructed by Dr. Paul Keim and Dr. Virginia Corvelli."

BASE COUNT 136 a 80 c 113 g 123 t  
ORIGIN

Alignment Scores:  
Pred. No.: 2.3e-67 Length: 452  
Score: 581.00 Matches: 112  
Percent Similarity: 97.50% Conservative: 5  
Best Local Similarity: 93.33% Mismatches: 3  
Query Match: 94.47% Indels: 0  
DB: 9 Gaps: 0

US-09-786-715-6 (1-120) x A1988470 (1-452)

QY 1 METALGLYSERERGLUGLYGINVALIISERCYSHSTHVRVALGLUTRPAAS 20  
DB 55 ATGGCAGGCTCCCGGAGGAGGAGCAAGCATTAAGCTGACACCCCTTGACGATGGAAC 114  
QY 21 ASPLINLEUGINLYSGLYASNGISERLYSLYLEULI VALASPHETHRALASER 40  
DB 115 GATCAACTCCAGAGGACCAACATCCAGAAACTATTTGTTGGATTTTACTGCTTGT 174  
QY 41 TRPCYSGIYPCYASRPHLEUALAPROPHLEUALAGLULEUALYLSYPHERH 60  
DB 175 TGGTGTGACACACCGCTTCATTTGACCATTTTGGCTGAGCTGCTAAAGTTTACA 234  
QY 61 SERVALIHPHEULYSVALASPLULEULYSSEVALSERGLASPTTRPAL 80  
DB 235 AGTGTGATCTCTTAAGGTCATGTGACGATTAAGAGTGTTCACAGACTGGCT 294  
QY 81 ILEGLUALMETPROTHRPHLEUALPHEVALYSGIUGLYTHRIEULASPLYVAL 100  
DB 295 ATTGAGGCTATGCGCCACTTTTGTGTGAAAGAGGAAACCTCTTAAGCAAGGTGG 354  
QY 101 GLVALYLSYASPLULEUGINLYSLIEGLINLYSHISVALASERVALASERVAL 120

	ORGANISM	TITLE	JOURNAL	COMMENT
REFERENCE	GLYCINE max			
AUTHORS	Eukaryotes: Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids ; eustoids I; Fabales; Fabaceae; Papilionoideae; Phaseoleae; Glycine.			
1 (bases 1 to 472)				
Shoemaker,R., Reim,P., Vodkin,L., Erpelding,J., Corvelli,V., Khanna,A., Bolla,B., Marra,M., Hillier,L., Kucaba,T., Martin,J., Beck,C., Wylie,T., Underwood,K., Stepien,M., Theisling,B., Allen,M., Bowers,Y., Person,B., Swaller,T., Gibbons,M., Page,D., Harvey,N., Schurk,R., Ritter,E., Korn,S., Shin,T., Jackson,Y., Cardenas,M., McCann,R., Waterston,R. and Wilson,R.				
Public Soybean EST Project				
Unpublished				
Contact: Shoemaker R/Public Soybean EST Project				
Public Soybean EST Project				
Washington University School of Medicine				
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA				
Tel: 314 286 1800				
Fax: 314 286 1810				
Email: est@watson.wustl.edu				
Predictive full length read				
vector to vector length is 622 This clone is available through:				
Reagen, Invitrogen Corp. 2130 South Memorial Parkway Huntsville, AL 35801 For further information call: (800)-533-4363 or contact via email: c@reagen.com				
High quality sequence stop: 441.				
Location/Qualifiers				
1..472				
/organism="Glycine max"				
/mol_type="mRNA"				
/db_xref="taxon:3847"				
/clone="GENOME SYSTEMS CLONE ID: Gm-cl051-1569"				
/tissue_type="floral metastemetic mRNA"				
/lab_host="DH10B"				
/clone_idb="Gm-cl051"				
/note="Vector: pluescript II SK+, Site.1: EcoRI, Site.2: XhoI; The CDNA library was constructed from floral metastemetic mRNA provided by Dr. Halina Knap of Clemson University. Complementary DNA was synthesized from mRNA using a primer consisting of a poly(dT) sequence with a blunt restriction site. EcoRI adapters were ligated to the blunt-ended cDNA fragments followed by XhoI digestion. The cDNA fragments were directionally cloned into the EcoRI-XhoI restriction site of the pluescript vector. The ligated cDNA fragments were transformed into DH10B host cells (GibcoBRL). This library was constructed in the laboratory of Dr. Randy Shoemaker."				
BASE COUNT	132 a 88 c 116 g 134 t	2 others		
ORIGIN				
Alignment Scores:				
Pred. NO.:	1.32e-67	Length:	472	
Score:	583.00	Matches:	118	
Percent Similarity:	99.16%	Conservative:	0	
Best Local Similarity:	99.16%	Mismatches:	1	
Query Match:	94.80%	Indels:	1	
DB:	10	Gaps:	0	
US-09-786-715-6 (1-120) x BE803058 (1-472)				
OY	2 AlaglyserseergluguglyglnValIIesercyshstprValgluglutrPasnasp 21			
Dd				
15 GCTGGCCTCATCGAAGAGGCACAGTGTATTAGTCGCCACACCCTTGAAATGAACGAT 74				
OY	22 GluleuglInylsglylaengluserLysylsleutllevalValasppherAlaseTrp 41			
Dd				
75 CCACTCCAGAAGGCCAACGAATCCMAAACAATCATTTGTGGAGATTACTGCTTCTGG 134				
OY	42 CysglyprocyarngphellealaprophenleualagluleualylslyspheThSer 61			
Dd				
135 TGTGGACCATGCGCTTTCATTGACCACTTCTGTGGCTAGCGGCTTAAGAAGTTACAGT 194				

REFERENCE	TITLE	JOURNAL	COMMENT
62	ValllephleuVasValAspValAspGluLeuYssSerValSerGlnAspTTPAlaIle 81		
195	GTCTATTCTCTAAAGGT-GATGTGGACCAATTTAAAGAGTCTTTCTCAAGATGGGCTATT 253		
82	GluAlaMetProThrPheValPheValLysGluGlyThrLeuLeuAspLysValValGly 101		
254	GAGCTATGGCCACTTTTGTGTGTTGTGTAAGAGGGAACGCTTCTGCAAAAGTGTGGGA 313		
102	AlaLysLysAspGluLeuGlnGlnLysIleGlnLysHisValAlaSerAlaSerAla 120		
314	GCAAAAGAAGATGAGTGCAGACAGAAATACAGAAACATGTGGCTTCAGCTAGTGCCT 370		
RESULT 9			
LOCUS			
DEFINITION			
ACCESSION			
VERSION			
KEYWORDS			
SOURCE			
ORGANISM			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
COMMENT			
FEATURES			
SOURCE			

## ORIGIN

## Alignment Scores:

Pred. No.: 2 186-71 Length: 445  
 Score: 611.00 Matches: 119  
 Percent Similarity: 99.17% Conservative: 0  
 Best Local Similarity: 99.17% Mismatches: 1  
 Query Match: 99.35% Indels: 0  
 DB: 13 Gaps: 0

US-09-786-715-6 (1-120) x BU578741 (1-445)

QY 1 MetAaGlySerSerGluGluGluValIleSerCysHisThrValGluGluTrpAsn 20  
 |||||||  
 Db 79 ATGCGTGGATCATCGAAGAGGACAAAGTCTATTGCGCCACCGTTGAAGATGCAAC 138  
 |||||||  
 QY 21 AspGlnLeuGlnGlyAsnGlnSerLysLysLeuIleValValAspPheThrAlaSer 40  
 |||||||  
 Db 139 GATCACTCCAGAAAGGCGACAGAACTCAATGTTGCGATTCTTACTGCTTCT 198  
 |||||||  
 QY 41 TrpCysGlyProCysArgPheIleAlaProPheLeuAlaGluAlaLysLysPheThr 60  
 |||||||  
 Db 199 TGGCTGTGACCATTCCTGTTCTTGCACCATTTCTGCTGAGCTGCTAAGAGTTCA 258  
 |||||||  
 QY 61 SerValIlePheLeuLysValAspValAspGluLeuLysSerValSerGlnAspTrpAla 80  
 |||||||  
 Db 259 AGTGTCAATATCTCAAGAGTGTGATGTGAGCAATTAAGAGTGTTCACAAGATTGGCT 318  
 |||||||  
 QY 81 IleGluAlaMetProThrPheValPheValLysGluGlyThrLeuLeuAspLysVal 100  
 |||||||  
 Db 319 ATTAGGCTATGCCACCTTTGTGTGTGAAGAGGAGAACGCTTGGACAAAGTGATG 378  
 |||||||  
 QY 101 GlnAlaLysLysAspGluLeuGlnGlnLysIleGlnLysHisValAlaSerAlaSerAla 120  
 |||||||  
 Db 379 GGACCAAGAAGAGATGAGCTGCACAGAAATACAGAAACATGTGGCTTCAAGTACGTGCT 438  
 |||||||

## RESULT 7

CA937588/c 572 bp mRNA linear EST 30-DEC-2002  
 LOCUS sav21g03.y1 Gm-c1048 Glycine max cDNA clone SOYBEAN CLOVE ID:  
 Gm-c1048-6078 5' similar to TR:Q43636 Q43636 THIOREDOXIN. ;, mRNA  
 DEFINITION sequence.

ACCESSION CA937588  
 VERSION CA937588.1 GI:27426068  
 KEYWORDS EST.  
 SOURCE Glycine max (soybean)  
 ORGANISM Glycine max

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids  
 ; eurosids I; Fabales; Fabaceae; Papilionoideae; Phaseoleae;  
 Glycine.

REFERENCE 1 (bases 1 to 572)  
 AUTHORS Shoemaker,R., Keim,P., Vodkin,L., Erpelting,J., Coryell,V., Khanna  
 A., Bolla,B., Marra,M., Hillier,L., Kucaba,T., Martin,J., Beck,C.,  
 Wylie,T., Underwood,K., Stepien,M., Theising,B., Allen,M., Bowers  
 ,T., Person,B., Swaller,T., Gibbons,M., Pape,D., Harvey,N., Schurk  
 ,R., Ritter,E., Kohn,S., Shih,T., Jackson,Y., Cardenas,M., McCann  
 ,R., Waterston,R. and Wilson,R.  
 TITLE Public Soybean EST Project  
 JOURNAL Unpublished  
 COMMENT Contact: Shoemaker R/Public Soybean EST Project  
 Public Soybean EST Project  
 Washington University School of Medicine  
 4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA  
 Tel: 314 286 1800  
 Fax: 314 286 1810  
 Email: est@watson.wustl.edu  
 This clone is available through: ResGen, Invitrogen Corp. 2130  
 South Memorial Parkway Hunttsville, AL 35801 For further information  
 call: (800)-533-4363 or contact: ccu@resgen.com web site:  
 www.resgen.com  
 Possible reversed clone: similarity on wrong strand  
 Seq primer: -40RP from Gibco

## FEATURES

High quality sequence stop: 435.  
 Location/Qualifiers  
 1..572  
 /organism="Glycine max"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:3847"  
 /clone="SOYBEAN CLOVE ID: Gm-c1048-6078"  
 /tissue\_type="whole seedlings of greenhouse grown plants"  
 /dev\_stage="1 week old"  
 /lab\_host="DH10B"  
 /clone\_id="Gm-c1048"  
 /note="Vector: pBluescript II SK+; site\_1: EcoRI; site\_2:  
 XhoI; The Clark NIL was constructed and seed was provided  
 by Dr. J. Specht, University of Nebraska (Shoemaker and  
 Specht, 1995). The cDNA library was constructed from mRNA  
 isolated from whole seedlings of 1 week old greenhouse  
 grown plants. Complementary DNA was synthesized from mRNA  
 using a primer consisting of a poly(dT) sequence with a  
 XhoI restriction site and a 3' anchor. EcoRI adapters were  
 ligated to the blunt-ended cDNA fragments followed by XhoI  
 digestion. The cDNA fragments were directionally cloned  
 into the EcoRI-XhoI restriction site of the pBluescript  
 vector. The ligated cDNA fragments were transformed into  
 DH10B host cells (GibcoBRL). The library was constructed  
 in cooperation with Dr. Paul Keim's laboratory at Northern  
 Arizona University."

BASE COUNT 174 a 125 c 102 g 171 t  
 ORIGIN

## Alignment Scores:

Pred. No.: 9 036-70 Length: 572  
 Score: 600.00 Matches: 117  
 Percent Similarity: 100.00% Conservative: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 97.56% Indels: 0  
 DB: 14 Gaps: 0

US-09-786-715-6 (1-120) x CA937588 (1-572)

QY 4 SerSerGluGluGlnGlnValIleSerCysHisThrValGluGluTrpAsnAspGlnLeu 23  
 |||||||  
 Db 572 TCATCGGAAGAAGGACAGCAAGTATTAGCTGCCACCCCTTAAGAGATGCAATCACTC 513  
 |||||||  
 QY 24 GlnLysGlnAsnGlnSerLysLysLeuIleValValAspPheThrAlaSerTrpCysGly 43  
 |||||||  
 Db 512 CAGAGGCGACAGCAATCCAGAAACATGTTGTGATTTTCTGCTTCTTGCTGCGA 453  
 |||||||  
 QY 44 ProCysArgPheIleAlaProPheLeuAlaGluLeuAlaLysLysPheThrSerValIle 63  
 |||||||  
 Db 452 CCATGCCCTTTCATTCGACCATCTTGGCTGAGCTGCTAAGAAAGTTCAACAAGTGCATA 393  
 |||||||  
 QY 64 PheLeuLysValAspValAspGluLeuLysSerValSerGlnAspTrpAlaIleGluAla 83  
 |||||||  
 Db 392 TTCCTTAAGGTGATGTGGAGCAATTAAGAGTGTTCGAAGATTGGCTATTGAGGCT 333  
 |||||||  
 QY 84 MetProThrPheValPheValLysGluGlyThrLeuLeuAspLysValValGlnAlaLys 103  
 |||||||  
 Db 332 ATGCCCACTTTTGTGTTTGAAGAAGAGCAAGCGTTCTGCAACAAGTGTGGCGCAAG 273  
 |||||||  
 QY 104 LysAspGluLeuGlnGlnLysIleGlnLysHisValAlaSerAlaSerAla 120  
 |||||||  
 Db 272 AAGATGAGCTGCAGCAGAAATACAGAAACATGTGCTTCACCTAGTGCT 222  
 |||||||

## RESULT 8

BE803058 472 bp mRNA linear EST 06-DEC-2001  
 LOCUS BE803058  
 DEFINITION Gm-c1051.y1 Gm-c1051 Glycine max cDNA clone GENOME SYSTEMS CLOVE ID:  
 Gm-c1051-1569 5' similar to TR:Q43636 Q43636 THIOREDOXIN. ;, mRNA  
 sequence.

ACCESSION BE803058  
 VERSION BE803058.1 GI:10234170  
 KEYWORDS EST.  
 SOURCE Glycine max (soybean)

TITLE  
JOURNAL  
COMMENT

, R., Waterston, R. and Wilson, R.  
Public Soybean EST Project  
Unpublished  
Contact: Shoemaker R/Public Soybean EST Project  
Public Soybean EST Project  
Washington University School of Medicine  
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA  
Tel: 314 286 1800  
Fax: 314 286 1810  
Email: est@waterston.wustl.edu  
This clone is available through: Resgen, Invitrogen Corp. 2130  
South Memorial Parkway Huntville, AL 35801 For further information  
call: (800)-533-4363 or contact: c@resgen.com web site:  
www.resgen.com  
Seq primer: -40RP from Gibco  
High quality sequence stop: 430.

## FEATURES

source

1.582  
Location/Qualifiers  
/organism="Glycine max"  
/mol\_type="mRNA"  
/db\_xref="taxon:3847"  
/clone="SOYBEAN CLONE ID: Gm-cl045-4100"  
/tissue\_type="Hypocotyl, 9-10 day old etiolated seedlings"  
/lab\_host="DH10B"  
/clone\_lib="Gm-cl045"  
/note="Vector: pBluescriptII SK+; Site:1: EcoRI; Site:2:  
XhoI; This cDNA library was constructed from mRNA isolated  
from etiolated hypocotyl tissue of 9-10 day old seedlings  
of the cultivar Williams 82. Complementary DNA was  
synthesized from mRNA using a primer consisting of a  
poly(dT) primer with a XhoI restriction site. EcoRI  
adapters were ligated to the blunt-ended cDNA fragments  
followed by digestion with EcoRI and XhoI. The cDNA  
fragments were directionally cloned into the EcoRI-XhoI  
restriction site of the pBluescript vector. The ligated  
cDNA fragments were transformed into DH10B host cells  
(Gibco BRL). This library was constructed by Dr. Randy  
Shoemaker."

BASE COUNT 166 a 105 c 135 g 176 t  
ORIGIN

## Alignment Scores:

Pred. No.: 9.08e-72 Length: 582  
Score: 615.00 Matches: 120  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 100.00% Indels: 0  
DB: 13 Gaps: 0

US-09-786-715-6 (1-120) x BQ630383 (1-582)

OY 1 MetAlaGlySerSerLunGluGlnValIleSerCyHisThrValGluGluTTPaen 20  
|||||  
DB 52 ATGGCGGCTCATCGAAGAGGACAAAGCATTAAGCTGCACACCTTGAGAGTGAAC 111  
|||||  
OY 21 AspGlnLeuGlnLysGlnGlnSerLysLeuIleValValAspPheThrAlaSer 40  
|||||  
DB 112 GATCAACTCCAGAGGCAACGATCCAGAACTCATGTTGTGATTTTACTGCTTCT 171  
|||||  
OY 41 TTPCyGlyProCySarPheIleAlaProPheLeuAlaGluLeuAlaLysLysPheThr 60  
|||||  
DB 172 TGGTGGAGCAATGCGCTTTCATTCGACATTCCTTGGCAGCTGCTAAGAGTTCACA 231  
|||||  
OY 61 SerValIlePheLeuLysValAspValAspGluLeuLysSerValSerGlnAspTyrAla 80  
|||||  
DB 232 AGTGCATATTCCTAAAGGTGATGAGAGCAATTAAGAGTCTCAAGATTGGCT 291  
|||||  
OY 81 IleGluAlaMetProThrPheValPheValLysGluGlyThrLeuLeuAspLysValAla 100  
|||||  
DB 292 ATTGAGGCTATGCGCCACTTTGTGTGGAAAGAGGCAACGCTTGGACAAAGTGTTG 351  
|||||  
OY 101 GYAlaLysLysAspGluLeuGlnGlnLysIleGlnLysHisValAlaSerAlaSerAla 120  
|||||

DB 352 GGAGCAAGAGAGATGAGCTGCAGACAAATACAGAAACATNGGCTTACGCTAGCTGCT 411

## RESULT 6

B0578741

LOCUS

DEFINITION

B0578741 445 bp mRNA linear EST 17-SEP-2002  
sa60h12.y1 Gm-cl074 Glycine max cDNA clone SOYBEAN CLONE ID:  
Gm-cl074-6911 5' similar to TR:043636 Q43636 THIOREDOXIN. ; mRNA  
sequence.

ACCESSION

B0578741

VERSION

B0578741.1 GI:23063968

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

Shoemaker, R., Keim, P., Vodkin, L., Erpellding, J., Coryell, V., Rhanna  
, A., Bolla, B., Marra, M., Hillier, L., Kucaba, T., Martin, J., Beck, C.,  
Wyllie, T., Underwood, K., Steptoe, M., Rheising, B., Allen, M., Bowers  
, Y., Peterson, B., Swaller, T., Gibbons, M., Pape, D., Harvey, N., Schurr  
, R., Ritter, E., Kohn, S., Shin, T., Jackson, Y., Cardenas, M., McCann  
, R., Waterston, R. and Wilson, R.  
Public Soybean EST Project  
Unpublished  
Contact: Shoemaker R/Public Soybean EST Project  
Public Soybean EST Project  
Washington University School of Medicine  
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA  
Tel: 314 286 1800  
Fax: 314 286 1810  
Email: est@waterston.wustl.edu  
This clone is available through: Resgen, Invitrogen Corp. 2130  
South Memorial Parkway Huntville, AL 35801 For further information  
call: (800)-533-4363 or contact: c@resgen.com web site:  
www.resgen.com  
Seq primer: -40RP from Gibco  
High quality sequence stop: 409.

## FEATURES

source

1.445  
Location/Qualifiers  
/organism="Glycine max"  
/mol\_type="mRNA"  
/db\_xref="taxon:3847"  
/clone="SOYBEAN CLONE ID: Gm-cl074-6911"  
/tissue\_type="seedlings induced for HR (hypersensitive  
response)"  
/dev\_stage="9-11 day old"  
/lab\_host="DH10B"  
/clone\_lib="Gm-cl074"  
/note="Vector: pBluescript II SK+; Site:1: EcoRI; Site:2:  
XhoI; The cDNA library was constructed from mRNA isolated  
from 9-11 day old seedlings that were induced for HR  
(hypersensitive response) by vacuum infiltrating plant  
tissue with *Pseudomonas syringae* pv. *glyciniae* carrying the  
avrB gene (Genetics 141:1597-1604). Plant tissue (expanded  
unifoliate leaves) was collected at 2, 4, 8, 12, 24, 36,  
and 53 hrs after inoculation and their mRNA pooled equally  
for cDNA construction. The library was prepared using the  
Stratagene pBluescript II SK(+) library construction kit.  
Complementary DNA was synthesized from mRNA using a primer  
consisting of a poly(dT) sequence with an XhoI restriction  
site. EcoRI adapters were ligated to the blunt-ended cDNA  
fragments followed by XhoI digestion. The cDNA insert is  
protected from XhoI digestion via methylation during first  
strand synthesis. The cDNA fragments were directionally  
cloned into the EcoRI-XhoI restriction site of the  
pBluescript vector. The ligated cDNA fragments were  
transformed into *E. coli* Electromax DH10B host cells. Plant  
care, inoculations, and library construction were  
performed by Steve Clough (Illa Vodkin lab, University of  
Illinois)."

BASE COUNT 130 a 81 c 118 g 116 t

Query Match: 100.00% Indels: 0  
 DB: 10 Gaps: 0  
 US-09-786-715-6 (1-120) x BG508580 (1-504)

OY 1 MetAlaGlySerSerGluGluGlyGlnValIleSerCysHisThrValGluGluTrpAsn 20  
 DB 37 ATGGCTGGCTCATCGAAGAGGAGGAGACAGTCTTACCTCCACACCGCTTGAGAAATGGAAC 96  
 OY 21 AspGlnLeuGlnIlySgIlyAsnGlnSerIlySylLeuIleValValAspPheThrAlaSer 40  
 DB 97 GATCAACTCCAGAAAGGACGACCAATCCAAAGAACTCATTTGTTGGATTTCATCTCTCT 156  
 OY 41 TrpCysGlyProCysArgPheIleAlaPheProPheLeuAlaGluLeuAlaIlySylSylPheThr 60  
 DB 157 TGGTGTGACCACTGCCGCTTTCATTCCTACCTCTTGGCTGAGCTGCTTACGAGTTCACA 216  
 OY 61 SerValIlePheLeuIlySylValAlaPylAspGluLeuIlySerValSerGlnAspTrpAla 80  
 DB 217 AGTGCATATTCCTTAAGCTGGATGGATGGACGAAATTAAGAGTTCATTCAGATGGGCT 276  
 OY 81 IleGluAlaMetProThrPheValAlaPheValIlySgIlyThrLeuLeuAspIlySylVal 100  
 DB 277 ATTGAGCTATGCCACCTTTGTGTGTGTGAAGAGGAGACGCTTCTGACAAACTGGTG 336  
 OY 101 GlyAlaIlySylAspGluLeuGlnGlnIlySylIleGlnIlySylHisValAlaSerAlaSerAla 120  
 DB 337 GGAGCAAGAAAGATGAGCTGCAGCAAGAAATACAGAAACATGTGGCTTCACGCTAGTCT 396

RESULT 4 B0630469 572 bp mRNA linear EST 21-OCT-2002  
 LOCUS B0630469  
 DEFINITION sag06f11.y1 Gm-cl045 glycine max cdna clone SOYBEAN CLONE ID: Gm-cl045-4125 5' similar to TR:Q43636 Q43636 THIOREDOXIN. ; mRNA sequence.  
 ACCESSION B0630469.1 GI:21678108  
 VERSION B0630469.1  
 KEYWORDS EST.  
 SOURCE Glycine max (soybean)  
 ORGANISM Glycine max

REFERENCE 1 (bases 1 to 572)  
 AUTHORS Shoemaker, R., Kelm, P., Vodkin, L., Erpelting, J., Coryell, V., Khanna, A., Bolla, B., Marra, M., Hillier, L., Kucaba, T., Martin, J., Beck, C., Wylie, T., Underwood, K., Steptoe, M., Theising, B., Allen, M., Bowers, Y., Person, B., Swaller, T., Gibbons, M., Pape, D., Harvey, N., Schurk, R., Ritter, E., Kohn, S., Shin, T., Jackson, Y., Cardenas, M., McCann, R., Waterston, R., and Wilson, R.  
 TITLE Public Soybean EST Project  
 JOURNAL Unpublished

COMMENT Contact: Shoemaker R./Public Soybean EST Project  
 Public Soybean EST Project  
 Washington University School of Medicine  
 4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA  
 Tel: 314 286 1800  
 Fax: 314 286 1810  
 Email: est@wustl.edu  
 This clone is available through: ResGen, Invitrogen Corp. 2130  
 South Memorial Parkway Huntville, AL 35801 For further information  
 call: (800)-533-4363 or contact: coudresgen.com web site:  
 www.resgen.com  
 Seq primer: ~40RP from Gibco  
 High quality sequence stop: 433.  
 Location/Qualifiers  
 1..572

FEATURES  
 source  
 /organism="Glycine max"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:3847"  
 /clone="SOYBEAN CLONE ID: Gm-cl045-4125"  
 /tissue\_type="Hypocotyl, 9-10 day old etiolated seedlings"

/lab\_host="DH10B"  
 /clone\_lib="Gm-cl045"  
 /note="Vector: pBluescriptII SK+; Site 1: EcoRI; Site 2: XhoI. This cDNA library was constructed from mRNA isolated from etiolated hypocotyl tissue of 9-10 day old seedlings of the cultivar Williams 82. Complementary DNA was synthesized from mRNA using a primer consisting of a poly(dT) primer with a XhoI restriction site. EcoRI adapters were ligated to the blunt-ended cDNA fragments followed by digestion with EcoRI and XhoI. The cDNA fragments were directionally cloned into the EcoRI-XhoI restriction site of the pBluescript vector. The ligated cDNA fragments were transformed into DH10B host cells (Gibco BRL). This library was constructed by Dr. Randy Shoemaker."

BASE COUNT 160 a 105 c 134 g 173 t  
 ORIGIN  
 Alignment Scores:  
 Pred. No.: 8,87e-72 Length: 572  
 Score: 615.00 Matches: 120  
 Percent Similarity: 100.00% Conservative: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 100.00% Indels: 0  
 DB: 13 Gaps: 0

US-09-786-715-6 (1-120) x B0630469 (1-572)

OY 1 MetAlaGlySerSerGluGluGlyGlnValIleSerCysHisThrValGluGluTrpAsn 20  
 DB 52 ATGGCTGGCTCATCGAAGAGGAGGAGACAGTCTTACCTCCACACCGCTTGAGAAATGGAAC 111  
 OY 21 AspGlnLeuGlnIlySgIlyAsnGlnSerIlySylLeuIleValValAspPheThrAlaSer 40  
 DB 112 GATCAACTCCAGAAAGGACGACCAATCCAAAGAACTCATTTGTTGGATTTCATCTCTCT 171  
 OY 41 TrpCysGlyProCysArgPheIleAlaPheProPheLeuAlaGluLeuAlaIlySylSylPheThr 60  
 DB 172 TGGTGTGACCACTGCCGCTTTCATTCCTACCTCTTGGCTGAGCTGCTTACGAGTTCACA 231  
 OY 61 SerValIlePheLeuIlySylValAlaPylAspGluLeuIlySerValSerGlnAspTrpAla 80  
 DB 232 AGTGCATATTCCTTAAGCTGGATGGACGCAAGAAATTAAGAGTTCATTCAGATGGGCT 291  
 OY 81 IleGluAlaMetProThrPheValAlaPheValIlySgIlyThrLeuLeuAspIlySylVal 100  
 DB 292 ATTGAGCTATGCCACCTTTGTGTGTGTGAAGAGGAGACGCTTCTGACAAACTGGTG 351  
 OY 101 GlyAlaIlySylAspGluLeuGlnGlnIlySylIleGlnIlySylHisValAlaSerAlaSerAla 120  
 DB 352 GGAGCAAGAAAGATGAGCTGCAGCAAGAAATACAGAAACATGTGGCTTCACGCTAGTCT 411

RESULT 5 B0630383 582 bp mRNA linear EST 21-OCT-2002  
 LOCUS B0630383  
 DEFINITION sag06f10.y1 Gm-cl045 glycine max cdna clone SOYBEAN CLONE ID: Gm-cl045-4100 5' similar to TR:Q43636 Q43636 THIOREDOXIN. ; mRNA sequence.  
 ACCESSION B0630383.1 GI:21678032  
 VERSION B0630383.1  
 KEYWORDS EST.  
 SOURCE Glycine max (soybean)  
 ORGANISM Glycine max

REFERENCE 1 (bases 1 to 582)  
 AUTHORS Shoemaker, R., Kelm, P., Vodkin, L., Erpelting, J., Coryell, V., Khanna, A., Bolla, B., Marra, M., Hillier, L., Kucaba, T., Martin, J., Beck, C., Wylie, T., Underwood, K., Steptoe, M., Theising, B., Allen, M., Bowers, Y., Person, B., Swaller, T., Gibbons, M., Pape, D., Harvey, N., Schurk, R., Ritter, E., Kohn, S., Shin, T., Jackson, Y., Cardenas, M., McCann, R., Waterston, R., and Wilson, R.

REFERENCE 1 (bases 1 to 582)  
 AUTHORS Shoemaker, R., Kelm, P., Vodkin, L., Erpelting, J., Coryell, V., Khanna, A., Bolla, B., Marra, M., Hillier, L., Kucaba, T., Martin, J., Beck, C., Wylie, T., Underwood, K., Steptoe, M., Theising, B., Allen, M., Bowers, Y., Person, B., Swaller, T., Gibbons, M., Pape, D., Harvey, N., Schurk, R., Ritter, E., Kohn, S., Shin, T., Jackson, Y., Cardenas, M., McCann, R., Waterston, R., and Wilson, R.

/db\_xref="taxon:3847"  
 /clone="SOYBEAN CLONE ID: Gm-c1072-3915"  
 /tissue\_type="seedlings induced for symptoms of SDS  
 (Sudden Death Syndrome) disease"  
 /dev\_stage="2-3 weeks old"  
 /lab\_host="DH10B"  
 /clone\_id="Gm-c1073"  
 /note="Vector: plasmid II SK+; Site\_1: EcoRI; Site\_2:  
 XhoI; The cDNA library was constructed from mRNA isolated  
 from 2-3 week old seedlings that were induced for symptoms  
 of SDS (Sudden Death Syndrome) disease by the  
 translocation of culture filtrate of *Fusarium solani* f.  
 sp. glycines (Plant Cell Report 18:375-380). Cultivar  
 Williams 82 is susceptible to the disease SDS. Plant  
 tissue (expanded leaves, folded leaves, and new shoots)  
 were collected at 1, 6, 24, and 48 hrs. after inoculation  
 and their mRNA pooled equally for cDNA construction. The  
 library was prepared using the Stratagene Bluescript II  
 SK(+) library construction kit. Complementary DNA was  
 synthesized from mRNA using a primer consisting of a  
 poly(dT) sequence with an XhoI restriction site. EcoRI  
 adaptors were ligated to the blunt-ended cDNA fragments  
 followed by XhoI digestion. The cDNA insert is protected  
 from XhoI digestion via methylation during first strand  
 synthesis. The cDNA fragments were directionally cloned  
 into the EcoRI-XhoI restriction site of the plasmid  
 vector. The ligated cDNA fragments were transformed into  
 E. coli Electromax DH10B host cells. Plants were inoculated  
 by Shuxian Li (Glen Hartman lab, University of Illinois).  
 Library was constructed by Reena Philip and Steve Clough  
 (Lila Vodkin lab, University of Illinois)."

BASE COUNT 138 a 88 c 120 g 149 t  
 ORIGIN

## Alignment Scores:

Pred. No.: 7.31e-72 Length: 495  
 Score: 615.00 Matches: 120  
 Percent Similarity: 100.00% Conservative: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 100.00% Indels: 0  
 DB: 13 Gaps: 0

US-09-786-715-6 (1-120) x B0627882 (1-495)

QY 1 MetAlaGlySerSerGluGluGlnValIleSerCysHisThrValGluGluTrpAsn 20  
 |||||  
 Db 35 ATGGCTGGCTCATCGAGAGGAGGACAGTCATTAGCTCCACACCGTTGACAGATGAC 94  
 |||||  
 QY 21 AspGlnLeuGlnLysGlnGluSerLysLysLeuIleValValAspPheThrAlaSer 40  
 |||||  
 Db 95 GATCAACTCCAGAGGAGGACAGATCCAGAACTCATTTGTCGATTTCAGCTTCT 154  
 |||||  
 QY 41 TrpCysGlyProCysArgPheIleAlaProPheLeuAlaGluLeuAlaLysLysPheThr 60  
 |||||  
 Db 155 TGGTGGGACCAAGCCGTTTCATTCACCTTTGGCTGAGCGCTAAGAGTTGACA 214  
 |||||  
 QY 61 SerValIlePheLeuLysValAspValAspGluLeuLysSerValSerGlnAspTrpAla 80  
 |||||  
 Db 215 AGTGTATATTTCTTAAGGTGATGAGACGATTAAGAGCTTTCTCAAGATTGGGCT 274  
 |||||  
 QY 81 IleGluAlaMetProThrPheValPheValLysGluGlnThrLeuLeuAspLysValVal 100  
 |||||  
 Db 275 ATTGAGGCTATGCCCACTTTTGTGTTGTAAGAGGAGGAGCGTTCTGACAAAGTGTG 334  
 |||||  
 QY 101 GlyAlaLysLysAspGlnLeuGlnGlnLysIleGlnLysHisValAlaSerAla 120  
 |||||  
 Db 335 GAGCAAGAGAGATGAGTGCAGCAAGAAATACAGAAACATGTGGCTTACCTAGTGTCT 394  
 |||||

## RESULT 3

LOCUS B508580 504 bp mRNA linear EST 28-NOV-2001  
 DEFINITION sac4f108.y1 Gm-c1072 glycine max cDNA clone GENOME SYSTEMS CLONE  
 ID: Gm-c1072-1048 5' similar to TR:043636 Q43636 THIORDOXIN. ;

ACCESSION B508580  
 VERSION B508580.1 GI:13479237  
 KEYWORDS EST.  
 SOURCE Glycine max (soybean)  
 ORGANISM Glycine max  
 Eukaryota: Virdiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids  
 ; eucosids I; Fabales; Fabaceae; Papilionoideae; Phaseoleae;  
 Glycine.

REFERENCE 1 (bases 1 to 504)

AUTHORS Shoemaker,R., Keim,P., Vodkin,L., Expelling,J., Coryell,V., Khanna  
 A., Bolla,B., Marra,M., Hillier,L., Kucaba,T., Martin,J., Beck,C.,  
 Wylie,T., Underwood,K., Stepien,M., Theising,B., Allen,M., Bowers  
 Y., Person,B., Swaller,T., Gibbons,M., Pape,D., Harvey,N., Schurk  
 R., Ritter,E., Kohn,S., Shin,T., Jackson,Y., Cardenas,M., McCann  
 R., Waterston,R. and Wilson,R.  
 Public Soybean EST Project  
 Unpublished

TITLE JOURNAL  
 COMMENT Contact: Shoemaker R/Public Soybean EST Project  
 Public Soybean EST Project  
 Washington University School of Medicine  
 4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA  
 Tel: 314 286 1800  
 Fax: 314 286 1810  
 Email: est@watson.wustl.edu

FEATURES  
 source This clone is available through: ResGen, Invitrogen Corp. 2130  
 South Memorial Parkway Huntsville, AL 35801 For further information  
 call: (800)-533-4363 or contact via email: ccu@resgen.com  
 High quality sequence stop: 439.  
 Location/Qualifiers

1..504

/organism="Glycine max"

/mol\_type="mRNA"

/db\_xref="taxon:3847"

/clone="GENOME SYSTEMS CLONE ID: Gm-c1072-1048"

/tissue\_type="seedlings induced for symptoms of SDS  
 (Sudden Death Syndrome) disease"

/dev\_stage="2-3 weeks old"

/lab\_host="DH10B"

/clone\_id="Gm-c1072"  
 /note="Vector: Bluescript II SK+; Site\_1: EcoRI; Site\_2:  
 XhoI; The cDNA library was constructed from mRNA isolated  
 from 2-3 week old seedlings that were induced for symptoms  
 of SDS (Sudden Death Syndrome) disease by the  
 translocation of culture filtrate of *Fusarium solani* f.  
 sp. glycines (Plant Cell Report 18:375-380). Cultivar PI  
 567374 is partially resistant to the disease SDS. Plant  
 tissue (expanded leaves, folded leaves, and new shoots)  
 were collected at 1, 6, 24, and 48 hrs. after inoculation  
 and their mRNA pooled equally for cDNA construction. The  
 library was prepared using the Stratagene Bluescript II  
 SK(+) library construction kit. Complementary DNA was  
 synthesized from mRNA using a primer consisting of a  
 poly(dT) sequence with an XhoI restriction site. EcoRI  
 adaptors were ligated to the blunt-ended cDNA fragments  
 followed by XhoI digestion. The cDNA insert is protected  
 from XhoI digestion via methylation during first strand  
 synthesis. The cDNA fragments were directionally cloned  
 into the EcoRI-XhoI restriction site of the plasmid  
 vector. The ligated cDNA fragments were transformed into  
 E. coli Electromax DH10B host cells. Plants were inoculated  
 by Shuxian Li (Glen Hartman lab, University of Illinois).  
 Library was constructed by Steve Clough (Lila Vodkin lab,  
 University of Illinois)."

BASE COUNT 143 a 88 c 123 g 150 t  
 ORIGIN

## Alignment Scores:

Pred. No.: 7.49e-72 Length: 504  
 Score: 615.00 Matches: 120  
 Percent Similarity: 100.00% Conservative: 0  
 Best Local Similarity: 100.00% Mismatches: 0

REFERENCE  
AUTHORS

, eucosids I; Fabales; Fabaceae; Papilionoideae; Phaseoleae;  
Glycine.  
1 (bases 1 to 459)  
Shoemaker, R., Keim, P., Vockin, L., Erpelidng, J., Corvett, V., Khanna  
, A., Bolla, B., Marra, M., Hillier, L., Kucaba, T., Martin, J., Beck, C.,  
Wyllie, T., Underwood, K., Steptoe, M., Theising, B., Allen, M., Bowers  
, Y., Person, B., Swaller, T., Gibbons, M., Pape, D., Harvey, N., Schurk  
, R., Ritter, E., Kohn, S., Shin, T., Jackson, Y., Cardenas, M., McCann  
, R., Waterston, R. and Wilson, R.  
Public Soybean EST Project

TITLE  
JOURNAL  
COMMENT

Unpublished  
Contact: Shoemaker R./Public Soybean EST Project  
Public Soybean EST Project  
Washington University School of Medicine  
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA  
Tel: 314 286 1800  
Fax: 314 286 1810  
Email: est@wustl.wustl.edu  
Putative full length read  
vector to vector length is 527 This clone is available through:  
ResGen, Invitrogen Corp. 2130 South Memorial Parkway Huntsville, AL  
35801 For further information call: (800)-533-4363 or contact via  
email: ccul@resgen.com  
Seq primer: -40RP from GIBCO  
High quality sequence stop: 424.  
Location/Qualifiers

FEATURES  
source

1..459  
/organism="Glycine max"  
/mol\_type="mRNA"  
/db\_xref="taxon:3847"  
/clone="GENOME SYSTEMS CLONE ID: Gm-cl081-878"  
/tissue\_type="Roots of 7 day old 'Brady' seedlings"  
/dev\_stage="7 days old"  
/lab\_host="DH10B"  
/clone\_lib="Gm-cl081"  
/note="Vector: p Bluescript II SK(+), Site.1: EcoRI, Site.2:  
XhoI; The mRNA was isolated from roots of 7 day old  
'Brady' seedlings that were mock infected 48 hours prior  
to harvest. Dr. Gary Stacey generously donated the  
tissue. The roots were flash-frozen in liquid nitrogen.  
Stratagene's cDNA Synthesis Kit (catalog number 200401)  
was used to synthesize the cDNA. First-strand synthesis  
was performed with 5-methyl dCTP, hence the ligated cDNA  
was hemimethylated. A modification of Stratagene's  
first-strand synthesis primer was used. An 'anchor'  
nucleotide (V-A, C, or G) was added to the 3' end of the  
primer (GAGAGAGAGAGAGAGAGACTGCTGAG(TT)18V) to anchor  
the primer at the 5' end of the poly(A) tract. After  
second-strand synthesis, the cDNA ends were filled in with  
cloned Pfu DNA, ligated to EcoRI adapters and subsequently  
phosphorylated. The cDNA was then precipitated and  
redissolved in sterile, RNase-, DNase-free water. The XhoI  
site within the first-strand synthesis primer was then  
restricted by digestion with XhoI from Promega (400/ul);  
all XhoI sites in the cDNA would be protected by their  
hemimethylated status. The cDNA constructs were  
size-fractionated with a 500bp cutoff, using Sephacryl  
S-500 High Resolution (Pharmacia Biotech) in a 2 mm  
diameter column and a bed volume of approximately 1ml. The  
column eluent was precipitated, redissolved, and ligated  
into Stratagene's pBluescript II XR predigested vector  
(pBluescript II SK(+)) vector that has been digested with  
EcoRI and XhoI, and phosphorylated by stratagene). This  
library was constructed in the laboratory of Dr. Paul Keim  
and Dr. Virginia H. Corvett at Northern Arizona  
University."

BASE COUNT  
ORIGIN

132 a 81 c 114 g 132 t

## Alignment Scores:

Pred. No.: 6.61e-72 Length: 459  
Score: 615.00 Matches: 120  
Percent Similarity: 100.00% Conservative: 0

Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 100.00% Indels: 0  
DB: 12 Gaps: 0

US-09-786-715-6 (1-120) x B1699372 (1-459)

QY 1 MetAlaGlySerSergIuGluGlyGlnValIleSerCysHisThrValGluGluTrpAsn 20  
DB 33 ATGCTGGCTCATCGTGAAGAGGACGACAGTCAATGAGTGGCCACCGCTGAAGATGCAAC 92  
QY 21 AspGlnIleuGlnIuGlyAsnGlnSerIuSylsIuIleValIaAspThrAlaSer 40  
DB 93 GATCAACTCCGACGAAGGGGCAACGAATCCAAAGCAATCATGCTGGATTTTACTGCTTCT 152  
QY 41 TrpCysGlyProCysAtsPheIleAlaProPheIuIaGluIleuAlaIuSylsPheThr 60  
DB 153 TGGTGTGGCCATCGCTTTCATTTGACACATTTCTGGCTGAGCTGCGCTCAAGAATTGACA 212  
QY 61 SerValIlePheIuSylsValaIaSpValaIaSpIuIeulysSerValaIaSerIaAspTrpAla 80  
DB 213 AGTGTCATATTCCTAAAGGTGAGTGGACGACATTAAGAGTGTTCGACAGATTGGGCT 272  
QY 81 IleGluAlaMetProThrPheValaIaPheValaIuSylsGluGlyThrIleuAspIuValaI 100  
DB 273 ATTGAGGCTATGCCACCTTTGTGTGTGAAGAGGCAACGCTTCTGGACAAAGTGGTG 332  
QY 101 GlyAlaIuSylsAspGluIuGlnGlnIuSylsIleGlnIuSylsValaIaSerAla 120  
DB 333 GGACCAAGAGAGATGACTGCTGACACGACAAATACAGAACTGTGGCTTCAGCTAGTGTCT 392

RESULT 2  
LOCUS B0627882  
DEFINITION sa655b02.y2 Gm-cl073 Glycine max cDNA clone SOYBEAN CLONE ID:  
Gm-cl073-3915 5' similar to TR:Q43636 Q43636 THIOREDOXIN.; mRNA  
sequence.

ACCESSION B0627882  
VERSION B0627882.1 GI:21675531  
KEYWORDS EST.  
SOURCE Glycine max (soybean)  
ORGANISM Glycine max (soybean)  
REFERENCE  
AUTHORS Shoemaker, R., Keim, P., Vockin, L., Erpelidng, J., Corvett, V., Khanna  
, A., Bolla, B., Marra, M., Hillier, L., Kucaba, T., Martin, J., Beck, C.,  
Wyllie, T., Underwood, K., Steptoe, M., Theising, B., Allen, M., Bowers  
, Y., Person, B., Swaller, T., Gibbons, M., Pape, D., Harvey, N., Schurk  
, R., Ritter, E., Kohn, S., Shin, T., Jackson, Y., Cardenas, M., McCann  
, R., Waterston, R. and Wilson, R.  
Public Soybean EST Project  
Unpublished  
Contact: Shoemaker R./Public Soybean EST Project  
Public Soybean EST Project  
Washington University School of Medicine  
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA  
Tel: 314 286 1800  
Fax: 314 286 1810  
Email: est@wustl.wustl.edu  
This clone is available through: ResGen, Invitrogen Corp. 2130  
South Memorial Parkway Huntsville, AL 35801 For further information  
call: (800)-533-4363 or contact: ccul@resgen.com web site:  
www.resgen.com  
Putative full length read  
vector to vector length is 605  
Seq primer: -40RP from GIBCO  
High quality sequence stop: 420.  
Location/Qualifiers

FEATURES  
source

1..495  
/organism="Glycine max"  
/mol\_type="mRNA"



GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - nucleic search, using frame\_plus\_p2n model

Run on: August 17, 2003, 19:20:27 : Search time 1622.86 Seconds  
(without alignments)  
1797.162 Million cell updates/sec

Title: US-09-786-715-6  
Perfect score: 615  
Sequence: 1 MAGSSEGGVISCHEVEMN.....GAKKDELOKIKKVASASA 120

Scoring table: BLOSUM62  
Xgapop 10.0, Xgapext 0.5  
Ygapop 10.0, Ygapext 0.5  
Fgapop 6.0, Fgapext 7.0  
Delop 6.0, Delext 7.0

Searched: 22781392 seqs, 12152238056 residues  
Total number of hits satisfying chosen parameters: 45562784

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Command line parameters:  
-MODEL=frame+ p2n model -DEV=yld  
-Q/cgq2.1/USFTO.spool/US09786715/runat\_11082003\_150514\_6048/app\_query.fasta.1.1052  
-DB=EST -QFMT=fastap -SUFFIX=1st -MINMATCH=0.1 -LOOPEL=0 -LOOPEXT=0  
-UNITS-bits -START=1 -END=-1 -MATRIX=blsum62 -TRANS=human40.cdi -LIST=45  
-DOCALIGN=200 -THR.SCORE=pct -THR.MAX=100 -THR.MIN=0 -ALIGN=15 -MODE=LOCAL  
-OUTFMT=psco -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=200000000  
-USRR=US09786715 @CNC3.1.1.7834 @runat\_11082003\_150514\_6048 -NCPU=6 -ICPU=3  
-NO\_MAP -LARGEIOUR -NEG.SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG  
-DEV.TIMEOUT=120 -WARN.TIMEOUT=30 -THREDS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6  
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : EST: \*  
1: em\_estba: \*  
2: em\_esthum: \*  
3: em\_estinu: \*  
4: em\_estnu: \*  
5: em\_estov: \*  
6: em\_estpl: \*  
7: em\_estro: \*  
8: em\_hic: \*  
9: gb\_est1: \*  
10: gb\_est2: \*  
11: gb\_hic: \*  
12: gb\_est3: \*  
13: gb\_est4: \*  
14: gb\_est5: \*  
15: em\_estfun: \*  
16: em\_estom: \*  
17: em\_gss\_hum: \*  
18: em\_gss\_luv: \*  
19: em\_gss\_pln: \*  
20: em\_gss\_vit: \*  
21: em\_gss\_fun: \*  
22: em\_gss\_mam: \*  
23: em\_gss\_mus: \*  
24: em\_gss\_pro: \*  
25: em\_gss\_tod: \*  
26: em\_gss\_ping: \*  
27: em\_gss\_vrl: \*  
28: gb\_gss1: \*

29: gb\_gss2: \*  
Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	615	100.0	459	12	B1699372
2	615	100.0	495	13	B0627882
3	615	100.0	504	10	BG508580
4	615	100.0	572	13	B0630459
5	615	100.0	582	13	B0630383
6	611	99.3	445	13	B0578741
7	600	97.6	445	14	CA937588
8	583	94.8	472	10	BE803058
9	582	94.6	522	13	B0786610
10	581	94.5	452	9	A1988470
11	581	94.5	466	9	AW569018
12	577	93.8	613	12	BG838879
13	566	92.0	627	14	CD401101
14	556	90.4	687	14	CD038084
15	533	86.7	570	14	CA901886
16	528	85.9	572	14	CA411141
17	515	83.7	632	14	CA896875
18	512	83.3	445	10	BE610138
19	512	83.3	512	14	CA794815
20	502	81.6	498	14	CA901885
21	501	81.5	311	12	B1094898
22	500	81.3	537	13	B0625213
23	500	81.3	655	14	CB291687
24	497	80.8	653	14	CB293759
25	495	80.5	517	14	CA795737
26	494	80.3	654	14	CB291686
27	490	79.7	394	14	CA901887
28	490	79.7	493	14	CB303550
29	490	79.7	546	10	BE053835
30	485	78.9	480	13	B0822062
31	485	78.9	555	13	B0875659
32	485	78.9	564	14	CA825633
33	485	78.9	587	13	B0863412
34	485	78.9	622	9	AW255457
35	483	78.5	579	14	CA825259
36	481	78.2	448	14	CA824794
37	480	78.0	581	14	CA824034
38	475.5	77.3	543	14	CB077020
39	470	76.4	634	9	AU300651
40	469	76.3	586	9	AW255195
41	468	76.1	316	12	B1892660
42	468	76.1	413	14	T24347
43	467	75.9	540	14	CB377001
44	467	75.9	602	10	BE054543
45	466	75.8	523	12	BM11010

## ALIGNMENTS

RESULT 1  
LOCUS B1699372  
DEFINITION B1699372 459 bp mRNA linear EST 29-NOV-2001  
sag36c07.y1 Gm-cl081 glycine max cDNA clone GENOME SYSTEMS CLONE  
ID: Gm-cl081-878 5' similar to TR:Q43636 Q43636 THIOREDOXIN. ;  
ACCESSION B1699372  
VERSION B1699372.1 GI:15662001  
KEYWORDS EST.  
SOURCE Glycine max (soybean)  
ORGANISM Glycine max  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids





US-09-786-715-6 (1-120) x US-10-091-841-5 (1-393)

QY 8 GlnValIleSerCysHisThrValGluGluTrpAsnAspGlnLeuGlnLysGlyAsn 27  
DB 61 GGGGAGGTGATCTCCGTCACAGCCTGAGCAGTGCACCATGCAGATCGAGAGGCCAAC 120  
QY 28 GluSerLysLysLeuIleValAlaAspPheThrAlaSerTrpCysGlyProCysArgPhe 47  
DB 121 GCGCCAGAGAGCTGTGTGTGATGATGATCTTCACTGCATCATGTGTGCGGACCATGCCGATT 180  
QY 48 IleAlaProPheLeuAlaGluLeuAlaLysLysPheThrSerValIlePheLeuLysVal 67  
DB 181 ATGCTCCCAATTTTGGCTGATCTCGCCAGAGATTGCCAGCTGCTGTTTCTCCTCAAGTTC 240  
QY 68 AspValAspGluLeuLysSerValSerGlnAspTrpAlaIleGluAlaMetProThrPhe 87  
DB 241 GACGTTGATGAACTGAAGCCCATCTGAGCAATTCAGGCTGAGGCCATGCCAACCTTTC 300  
QY 88 ValPheValLysGluGlyThrLeuLeuAspLysValAlaGlyAlaLysLysAspGluLeu 107  
DB 301 CTGTTTCATGAGGAGAGAGAGCTCAAGACAGAGGTTGTCGAGCTATCAAGAGAGAGCTG 360  
QY 108 GlnGlnLysIleGlnLysHisValAlaSer 117  
DB 361 ACGACCAAGGTTGGCTCCACCGCGCTGCC 390

RESULT 13  
US-10-091-841-3  
Sequence 3, Application US/10091841  
Publication No. US20030150010A1  
GENERAL INFORMATION:  
APPLICANT: Cho, Myeong-Je  
APPLICANT: Del Val, Greg  
APPLICANT: Caillaud, Maxime  
APPLICANT: Lemaux, Peggy G.  
APPLICANT: Buchanan, Bob B.  
TITLE OF INVENTION: Barley Gene for Thioredoxin and  
TITLE OF INVENTION: NADP-Thioredoxin Reductase  
FILE REFERENCE: 2001-0701.30  
CURRENT FILING DATE: 2002-03-05  
PRIOR FILING DATE: 09/540, 014  
PRIOR FILING DATE: 2000-03-31  
PRIOR APPLICATION NUMBER: US 60/127,198  
PRIOR FILING DATE: 1999-03-31  
PRIOR APPLICATION NUMBER: US 60/169,162  
PRIOR FILING DATE: 1999-12-06  
PRIOR APPLICATION NUMBER: US 60/177,740  
PRIOR FILING DATE: 2000-01-21  
PRIOR APPLICATION NUMBER: US 60/177,739  
PRIOR FILING DATE: 2000-01-21  
NUMBER OF SEQ ID NOS: 51  
SOFTWARE: FastSeq for Windows Version 4.0  
SEQ ID NO 3  
LENGTH: 382  
TYPE: DNA  
ORGANISM: Artificial Sequence  
FEATURE:  
OTHER INFORMATION: wheat thioredoxin h cDNA  
US-10-091-841-3

Alignment Scores:  
Pred. No.: 5.38e-45 Length: 382  
Score: 367.00 Matches: 63  
Percent Similarity: 81.65% Conservative: 26  
Best Local Similarity: 57.80% Mismatches: 20  
Query Match: 59.67% Indels: 0  
Gaps: 0

US-09-786-715-6 (1-120) x US-10-091-841-3 (1-382)

QY 8 GlnValIleSerCysHisThrValGluGluTrpAsnAspGlnLeuGlnLysGlyAsn 27

DB 52 GGGGAGGTGATCTCCGTCACAGCCTGAGCAGTGCACCATGCAGATCGAGAGGCCAAC 111  
QY 28 GluSerLysLysLeuIleValAlaAspPheThrAlaSerTrpCysGlyProCysArgPhe 47  
DB 112 GCGCCAGAGAGCTGTGTGTGATGATGATCTTCACTGCATCATGTGTGCGGACCATGCCGATT 171  
QY 48 IleAlaProPheLeuAlaGluLeuAlaLysLysPheThrSerValIlePheLeuLysVal 67  
DB 172 ATGCTCCCAATTTTGGCTGATCTCGCCAGAGATTGCCAGCTGCTGTTTCTCCTCAAGTTC 231  
QY 68 AspValAspGluLeuLysSerValSerGlnAspTrpAlaIleGluAlaMetProThrPhe 87  
DB 232 GACGTTGATGAACTGAAGCCCATCTGAGCAATTCAGGCTGAGGCCATGCCAACCTTTC 291  
QY 88 ValPheValLysGluGlyThrLeuLeuAspLysValAlaGlyAlaLysLysAspGluLeu 107  
DB 292 CTGTTTCATGAGGAGAGAGATCTCAAGACAGAGGTTGTCGAGCTATCAAGAGAGAACTG 351  
QY 108 GlnGlnLysIleGlnLysHisValAla 116  
DB 352 ACGACCAAGGTTGGCTACACCGCGGCC 378

RESULT 14  
US-10-194-885-8  
Sequence 8, Application US/10194885  
Publication No. US20030135878A1  
GENERAL INFORMATION:  
APPLICANT: Wong, J. H.  
APPLICANT: Cho, Myeong-Je  
APPLICANT: Lemaux, Peggy G.  
APPLICANT: Buchanan, Bob  
TITLE OF INVENTION: TRANSGENIC PLANTS WITH ELEVATED  
TITLE OF INVENTION: THIOREDOXIN LEVELS  
FILE REFERENCE: 416272000800  
CURRENT FILING DATE: 2002-07-12  
PRIOR FILING DATE: 2002-07-12  
PRIOR APPLICATION NUMBER: US/10/194,885  
PRIOR FILING DATE: 2001-07-19  
PRIOR APPLICATION NUMBER: 60/307,006  
PRIOR FILING DATE: 09/538,864  
PRIOR FILING DATE: 2000-03-29  
PRIOR APPLICATION NUMBER: 60/126,736  
PRIOR FILING DATE: 1999-03-29  
NUMBER OF SEQ ID NOS: 55  
SOFTWARE: FastSeq for Windows Version 4.0  
SEQ ID NO 8  
LENGTH: 369  
TYPE: DNA  
ORGANISM: Barley  
FEATURE:  
NAME/KEY: CDS  
LOCATION: (1)...(369)  
US-10-194-885-8

Alignment Scores:  
Pred. No.: 3.97e-44 Length: 369  
Score: 361.00 Matches: 62  
Percent Similarity: 81.65% Conservative: 27  
Best Local Similarity: 56.88% Mismatches: 20  
Query Match: 58.70% Indels: 0  
Gaps: 0

US-09-786-715-6 (1-120) x US-10-194-885-8 (1-369)

QY 9 GlnValIleSerCysHisThrValGluGluTrpAsnAspGlnLeuGlnLysGlyAsn 28  
DB 37 GAGGTGATCTCGTCACAGCCTGAGCAGTGCACCATGCAGATCGAGAGGCCAACACC 96  
QY 29 SerLysLysLeuIleValAlaAspPheThrAlaSerTrpCysGlyProCysArgPhe 48  
DB 97 GCCAAGAGAGCTGTGTGTGATGATGATCTTCACTGCATCATGTGTGCGGACCATGCCGATCAG 156  
QY 49 AlaProPheLeuAlaGluLeuAlaLysLysPheThrSerValIlePheLeuLysValAsp 68

SEQ ID NO 1736  
 LENGTH: 357  
 TYPE: DNA  
 ORGANISM: Arabidopsis thaliana  
 US-09-938-842A-1736

Alignment Scores:

Pred. No.: 1.9e-46 Length: 357  
 Score: 376.50 Matches: 67  
 Percent Similarity: 80.87% Conservative: 26  
 Best Local Similarity: 58.26% Mismatches: 19  
 Query Match: 61.22% Indels: 3  
 Gaps: 1

US-09-786-715-6 (1-120) x US-09-938-842A-1736 (1-357)

QY 7 GluGluGlnValIleSerCysHisThrValGluGluTrpAsnAspGlnLeuGlnLysGly 26  
 Db 10 GAAGAGAGAAGTATTCGCTTGCCACACCGTTGAGATTGGACCGAAGCTCAAGCCGCC 69  
 QY 27 AsnGluSerLysLysLeuIleValValAspPheThrAlaSerTPrcysGlyProCysArg 46  
 Db 70 AACGAATCCAAAGAACTGATTGTGATAGACTTCACTGCAACATGCTGCCACCTTGCGGT 129  
 QY 47 PheIleAlaProPheLeuAlaGluLeuAlaLysLysPheThrSerValIlePheLeuLys 66  
 Db 130 TTCATTGCAACCGCTCTTGTGACTTGCACCAAGACCTCGACGTAAGCTTCTTCAAG 189  
 QY 67 ValAspValAspGluLeuLysSerValSerGlnAspTrpAlaIleGluAlaMetProThr 86  
 Db 190 GTGCATGTTGACGAATGGAACACCTGTTGCTGAGAGATTAAAGTTCAAGGCAATCCAAAG 249  
 QY 87 PheValPheValLysGluGlyThrLeuLeuAspLysValValGlyAlaLysAspGlu 106  
 Db 250 TTATATCTTCATGAAGAAGAGAGATCAAGAGACTGTGTTGCTGCTGAAGAAGAA 309  
 QY 107 LeuGlnGlnLysIleGlnLysHis-----ValAlaSerAla 118  
 Db 310 ATCATTGCCAATCTCGAAGACAGACAGACTTGTCTGCTGCT 354

RESULT 11

US-10-323-362-1  
 Sequence 1, Application US/10323362  
 Publication No. US20030113786A1  
 GENERAL INFORMATION:  
 APPLICANT: Kurnik, Betsy  
 APPLICANT: Davis, Keith  
 APPLICANT: Zayed, Adel  
 APPLICANT: Ascenzi, Robert  
 APPLICANT: Harper, Angel  
 APPLICANT: Boyes, Douglas  
 APPLICANT: Mulpuri, Rao  
 APPLICANT: Hoffman, Neil  
 APPLICANT: Kjemtrup, Susanne  
 APPLICANT: Moessner, Jeffrey  
 APPLICANT: Gorlach, Jorn  
 APPLICANT: Hamilton, Carol  
 TITLE OF INVENTION: METHODS FOR THE IDENTIFICATION OF INHIBITORS OF THIOREDOXIN EXPRESSION  
 FILE REFERENCE: 2133US  
 CURRENT APPLICATION NUMBER: US/10/323,362  
 CURRENT FILING DATE: 2002-12-18  
 NUMBER OF SEQ ID NOS: 4  
 SOFTWARE: PatentIn version 3.1  
 SEQ ID NO 1  
 LENGTH: 357  
 TYPE: DNA  
 ORGANISM: Arabidopsis thaliana  
 US-10-323-362-1

Alignment Scores:  
 Pred. No.: 1.9e-46 Length: 357  
 Score: 376.50 Matches: 67

Percent Similarity: 80.87% Conservative: 26  
 Best Local Similarity: 58.26% Mismatches: 19  
 Query Match: 61.22% Indels: 3  
 Gaps: 1

US-09-786-715-6 (1-120) x US-10-323-362-1 (1-357)

QY 7 GluGluGlnValIleSerCysHisThrValGluGluTrpAsnAspGlnLeuGlnLysGly 26  
 Db 10 GAAGAGAGAAGTATTCGCTTGCCACACCGTTGAGATTGGACCGAAGCTCAAGCCGCC 69  
 QY 27 AsnGluSerLysLysLeuIleValValAspPheThrAlaSerTPrcysGlyProCysArg 46  
 Db 70 AACGAATCCAAAGAACTGATTGTGATAGACTTCACTGCAACATGCTGCCACCTTGCGGT 129  
 QY 47 PheIleAlaProPheLeuAlaGluLeuAlaLysLysPheThrSerValIlePheLeuLys 66  
 Db 130 TTCATTGCAACCGCTCTTGTGACTTGCACCAAGACCTCGACGTAAGCTTCTTCAAG 189  
 QY 67 ValAspValAspGluLeuLysSerValSerGlnAspTrpAlaIleGluAlaMetProThr 86  
 Db 190 GTGCATGTTGACGAATGGAACACCTGTTGCTGAGAGATTAAAGTTCAAGGCAATCCAAAG 249  
 QY 87 PheValPheValLysGluGlyThrLeuLeuAspLysValValGlyAlaLysAspGlu 106  
 Db 250 TTATATCTTCATGAAGAAGAGATCAAGAGACTGTGTTGCTGCTGAAGAAGAA 309  
 QY 107 LeuGlnGlnLysIleGlnLysHis-----ValAlaSerAla 118  
 Db 310 ATCATTGCCAATCTCGAAGACAGACAGACTTGTCTGCTGCT 354

RESULT 12

US-10-091-841-5  
 Sequence 5, Application US/10091841  
 Publication No. US20030150010A1  
 GENERAL INFORMATION:  
 APPLICANT: Cho, Myeong-Je  
 APPLICANT: Del Val, Greg  
 APPLICANT: Calliau, Maxime  
 APPLICANT: Lemaux, Peggy G.  
 APPLICANT: Buchanan, Bob B.  
 TITLE OF INVENTION: Barley Gene for Thioredoxin and  
 FILE REFERENCE: NADP-Thioredoxin Reductase  
 CURRENT APPLICATION NUMBER: US/10/091,841  
 CURRENT FILING DATE: 2002-03-05  
 PRIOR APPLICATION NUMBER: 09/540,014  
 PRIOR FILING DATE: 2000-03-31  
 PRIOR APPLICATION NUMBER: US 60/127,198  
 PRIOR FILING DATE: 1999-03-31  
 PRIOR APPLICATION NUMBER: US 60/169,162  
 PRIOR FILING DATE: 1999-12-06  
 PRIOR APPLICATION NUMBER: US 60/177,740  
 PRIOR FILING DATE: 2000-01-21  
 PRIOR APPLICATION NUMBER: US 60/177,739  
 PRIOR FILING DATE: 2000-01-21  
 NUMBER OF SEQ ID NOS: 51  
 SOFTWARE: FastSeq for Windows Version 4.0  
 SEQ ID NO 5  
 LENGTH: 393  
 TYPE: DNA  
 ORGANISM: Artificial Sequence  
 FEATURE:  
 OTHER INFORMATION: wheat thioredoxin h cDNA  
 US-10-091-841-5

Alignment Scores:  
 Pred. No.: 3.99e-45 Length: 393  
 Score: 368.00 Matches: 63  
 Percent Similarity: 81.82% Conservative: 27  
 Best Local Similarity: 57.27% Mismatches: 20  
 Query Match: 59.84% Indels: 0  
 Gaps: 0

Db 369 GACCTTCAAGCCACCATA 386

RESULT 8  
US-10-349-782-12  
; Sequence 12, Application US/10349782  
; Publication No. US20030143618A1  
; GENERAL INFORMATION:  
; APPLICANT: Yves Hatzfeld  
; APPLICANT: Valerie Marie-No. US20030143618A1lle Frankard  
; APPLICANT: Anne-Marie Droual  
; TITLE OF INVENTION: Method for easy cloning and selection of chimeric DNA molecules  
; FILE REFERENCE: 1187-15  
; CURRENT APPLICATION NUMBER: US/10/349,782  
; CURRENT FILING DATE: 2003-01-23  
; PRIOR APPLICATION NUMBER: EP 02075373.7  
; PRIOR FILING DATE: 2002-01-23  
; NUMBER OF SEQ ID NOS: 12  
; SOFTWARE: PatentIn version 3.1  
; LENGTH: 540  
; TYPE: DNA  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Modified thioredoxin of *Oryza sativa* in vector pDONR201  
US-10-349-782-12

Alignment Scores:  
Pred. No.: 9,71e-48 Length: 540  
Score: 387.00 Matches: 72  
Percent Similarity: 77.69% Conservative: 22  
Best Local Similarity: 59.50% Mismatches: 23  
Query Match: 62.93% Indels: 4  
DB: 12 Gaps: 1

US-09-786-715-6 (1-120) x US-10-349-782-12 (1-540)

QY 4 SerSerGluGluGlyGlnValIleSerCysHisThrValGluGluTrpAsnAspGlnLeu 23  
Db 61 GCGCGCAGAGAGGAGGAGTGTGATCGGCTCCACACAGAGAGAGTTCAGCCCGCAGAG 120

QY 24 GlnLysGlyAsnGluSerLysLysLeuIleValAlaAspPheThrAlaSerTrpCysGly 43  
Db 121 ACCAAGGCCAGAGAGCGCCGCAAGAGTGCATATGACTTCACCTGCTGCTGTCGGA 180

QY 44 ProCysArgPheIleAlaProPheLeuAlaGluLeuAlaLysLysPheThrValIle 63  
Db 181 CCGTCCCGCTTATCGCCCGCAGTGTTCGCTGAATACGCCAAAAGTTCCTGCTGCTGTC 240

QY 64 PheLeuLysValAspValAspGluLeuLysSerValSerGlnAspTrpAlaIleGluAla 83  
Db 241 TTCCCTGAAGAGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 300

QY 84 MetProThrPheValAlaPheValLysGluGlyThrLeuLeuAspLysValAlaGlyAlaLys 103  
Db 301 ATGCCGACCTTCTTATTCATCAAGATGATGATGATGATGATGATGATGATGATGATGATG 360

QY 104 LysAspGluLeuGlnGlnLysIleGlnLysHisVal-----AlaSerAlaSer 119  
Db 361 AAGGATGACCTTCAGAACACACATCGTGAAGCACGTCGCTGCCACTGCTGCTGCTGCT 420

QY 120 Ala 120  
Db 421 GCC 423

RESULT 9  
US-09-938-842A-164  
; Sequence 164, Application US/09938842A  
; Patent No. US20020160378A1  
; GENERAL INFORMATION:  
; APPLICANT: Harper, Jeff  
; APPLICANT: Kreps, Joel  
; APPLICANT: Wang, Xun  
; APPLICANT: Zhu, Tong

; TITLE OF INVENTION: STRESS-REGULATED GENES OF PLANTS, TRANSGENIC PLANTS CONTAINING  
; TITLE OF INVENTION: SAME, AND METHODS OF USE  
; FILE REFERENCE: SRIPI300-3  
; CURRENT APPLICATION NUMBER: US/09/938,842A  
; CURRENT FILING DATE: 2001-08-24  
; PRIOR APPLICATION NUMBER: US 60/227,866  
; PRIOR FILING DATE: 2000-08-24  
; PRIOR APPLICATION NUMBER: US 60/264,647  
; PRIOR FILING DATE: 2001-01-16  
; PRIOR APPLICATION NUMBER: US 60/300,111  
; NUMBER OF SEQ ID NOS: 5379  
; SEQ ID NO 164  
; LENGTH: 360  
; TYPE: DNA  
; ORGANISM: *Arabidopsis thaliana*  
US-09-938-842A-164

Alignment Scores:  
Pred. No.: 3,48e-47 Length: 360  
Score: 381.50 Matches: 72  
Percent Similarity: 77.12% Conservative: 19  
Best Local Similarity: 61.02% Mismatches: 26  
Query Match: 62.03% Indels: 1  
DB: 10 Gaps: 1

US-09-786-715-6 (1-120) x US-09-938-842A-164 (1-360)

QY 4 SerSerGluGluGlyGlnValIleSerCysHisThrValGluGluTrpAsnAspGlnLeu 23  
Db 4 GCGCGCAGAGAGGAGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 63

QY 24 GlnLysGlyAsnGluSerLysLysLeuIleValAlaAspPheThrAlaSerTrpCysGly 43  
Db 64 GATTAAGCCAAAGAACCAACAGAGTGTGATGATGATGATGATGATGATGATGATGATGATG 123

QY 44 ProCysArgPheIleAlaProPheLeuAlaGluLeuAlaLysLysPhe--ThrSerVal 62  
Db 124 CCATGCCGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 183

QY 63 IlePheLeuLysValAspValAspGluLeuLysSerValSerGlnAspTrpAlaIleGlu 82  
Db 184 ATCTTCTCAAGAGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 243

QY 83 AlaMetProThrPheValAlaPheValLysGluGlyThrLeuLeuAspLysValAlaGlyAla 102  
Db 244 GCATGCCAACCTTGTGTTCAATTAAGCCGCGGAGTGTGATGATGATGATGATGATGATG 303

QY 103 LysLysAspGluLeuGlnGlnLysIleGlnLysHisValAlaSerAlaSerAla 120  
Db 304 AATAAGAGATCTTCAGCGCAAAATATGAGAACATACCTGCTTACCAACTGCG 357

RESULT 10  
US-09-938-842A-1736  
; Sequence 1736, Application US/09938842A  
; Patent No. US20020160378A1  
; GENERAL INFORMATION:  
; APPLICANT: Harper, Jeff  
; APPLICANT: Kreps, Joel  
; APPLICANT: Wang, Xun  
; APPLICANT: Zhu, Tong  
; TITLE OF INVENTION: STRESS-REGULATED GENES OF PLANTS, TRANSGENIC PLANTS CONTAINING  
; TITLE OF INVENTION: SAME, AND METHODS OF USE  
; FILE REFERENCE: SRIPI300-3  
; CURRENT APPLICATION NUMBER: US/09/938,842A  
; CURRENT FILING DATE: 2001-08-24  
; PRIOR APPLICATION NUMBER: US 60/227,866  
; PRIOR FILING DATE: 2000-08-24  
; PRIOR APPLICATION NUMBER: US 60/264,647  
; PRIOR FILING DATE: 2001-01-16  
; PRIOR APPLICATION NUMBER: US 60/300,111  
; PRIOR FILING DATE: 2001-06-22  
; NUMBER OF SEQ ID NOS: 5379

Alignment Scores:  
Pred. No.: 1,71e-55 Length: 402  
Score: 438.00 Matches: 81  
Percent Similarity: 84.21% Conservative: 15  
Best Local Similarity: 71.05% Mismatches: 18  
Query Match: 71.22% Indels: 0  
DB: 10 Gaps: 0

US-09-786-715-6 (1-120) x US-09-878-574-3723 (1-402)

QY 5 SerGluGluGlnValIleSerCysHisThrValGluGluTrpAsnAspGlnLeuGln 24  
DB 46 GCGGAAGAGGACAAAGATGGGTGTCACAGCGGTGAGAGGAGGAACATCTCAAG 105  
QY 25 LysGluAsnGluSerLysLysLeuValValAspPheThrAlaSerTrpCysGlyPro 44  
DB 106 AAGGAGAGAGGACCAAGAAATGATTTGTTGATTTTCTGCTTCTTGTGGGTCGA 165  
QY 45 CysArgPheIleAlaProPheLeuAlaGluLeuAlaLysLysPheThrSerValIlePhe 64  
DB 166 TGGCGTTTCATTCGCCCAATTCCTTGACAGAGTTGCTAAAGATTGCCGAAATGCACCTTC 225  
QY 65 LeuLysValAspValAspGluLeuLysSerValSerGlnAspTrpAlaIleGluAlaMet 84  
DB 226 CTCACGTCGATGTGATGATGATGAGACTGTTCCACAGCAGTGGGAAATTTGAGGCTATG 285  
QY 85 ProThrPheValPheValLysGluGlyThrLeuLeuAspLysValValGlyAlaLys 104  
DB 286 CCACCTTCCTCTTCTTGAAGAGAGCAGCAGCTGTGACAGAGTTGGTGGCCACACAA 345  
QY 105 AspGluLeuGlnLysLysIleGlnLysHisValAlaSerAla 118  
DB 346 GAGAGCTGCATTTGACCTACGACAGACATGCGCAACTGCC 387

## RESULT 6

US-09-878-574-104  
; Sequence 104, Application US/09878574  
; Patent No. US20020110548A1  
; GENERAL INFORMATION:  
; APPLICANT: Byrum, Joseph R.  
; APPLICANT: La Rosa, Thomas J.  
; APPLICANT: Thompson, Michael D.  
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated with  
; FILE REFERENCE: 38-21(15401)B  
; CURRENT APPLICATION NUMBER: US/09/878,574  
; CURRENT FILING DATE: 2001-12-21  
; PRIOR APPLICATION NUMBER: 09/333,535  
; PRIOR FILING DATE: 1999-06-14  
; NUMBER OF SEQ ID NOS: 15775  
; SEQ ID NO 104  
; LENGTH: 392  
; TYPE: DNA  
; ORGANISM: Glycine max  
; OTHER INFORMATION: Clone ID: LIB3028-056-Q1-B1-B3  
US-09-878-574-104

## Alignment Scores:

Pred. No.: 1,96e-53 Length: 392  
Score: 424.00 Matches: 77  
Percent Similarity: 86.21% Conservative: 23  
Best Local Similarity: 66.38% Mismatches: 16  
Query Match: 68.94% Indels: 0  
DB: 10 Gaps: 0

US-09-786-715-6 (1-120) x US-09-878-574-104 (1-392)

QY 5 SerGluGluGlnValIleSerCysHisThrValGluGluTrpAsnAspGlnLeuGln 24  
DB 9 GCGAAGAGGAGTCAGCTTGGCGCTCCACACCGTTCGATGAGCAACAGCAGCAACTGCAG 68  
QY 25 LysGluAsnGluSerLysLysLeuValValAspPheThrAlaSerTrpCysGlyPro 44

DB 69 AATGAAAGACCTCCCAAAATGATTTGTTGATTTTACTGCTTCCTGTTGTCGA 128  
QY 45 CysArgPheIleAlaProPheLeuAlaGluLeuAlaLysLysPheThrSerValIlePhe 64  
DB 129 TGGCGTTTATTGCCCCAGTTCTGGCAGAAATTCCTAGGCACTCTCAAGTATCTTC 188  
QY 65 LeuLysValAspValAspGluLeuLysSerValSerGlnAspTrpAlaIleGluAlaMet 84  
DB 189 CTCACGTCGATGTGATGAGAGCGCTGTTGCTGAGGAATATTCATTTAGGCGCATG 248  
QY 85 ProThrPheValPheValLysGluGlyThrLeuLeuAspLysValValGlyAlaLys 104  
DB 249 CCACCTTCCTCTTCTTGAAGAGAGTGCAGATGCTGTGATAAGTTGTTGGTCTTAAGAG 308  
QY 105 AspGluLeuGlnLysLysIleGlnLysHisValAlaSerAla 120  
DB 309 GAGAGCTGCACACTCACCATAGCAGACATGATCTGCTGCTCT 356

## RESULT 7

US-09-878-574-1462  
; Sequence 1462, Application US/09878574  
; Patent No. US20020110548A1  
; GENERAL INFORMATION:  
; APPLICANT: Byrum, Joseph R.  
; APPLICANT: La Rosa, Thomas J.  
; APPLICANT: Thompson, Michael D.  
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated with  
; FILE REFERENCE: 38-21(15401)B  
; CURRENT APPLICATION NUMBER: US/09/878,574  
; CURRENT FILING DATE: 2001-12-21  
; PRIOR APPLICATION NUMBER: 09/333,535  
; PRIOR FILING DATE: 1999-06-14  
; NUMBER OF SEQ ID NOS: 15775  
; SEQ ID NO 1462  
; LENGTH: 390  
; TYPE: DNA  
; ORGANISM: Glycine max  
; OTHER INFORMATION: Clone ID: LIB3028-037-Q1-B1-F11  
US-09-878-574-1462

## Alignment Scores:

Pred. No.: 1,53e-48 Length: 390  
Score: 391.00 Matches: 72  
Percent Similarity: 84.91% Conservative: 18  
Best Local Similarity: 67.92% Mismatches: 16  
Query Match: 63.58% Indels: 0  
DB: 10 Gaps: 0

US-09-786-715-6 (1-120) x US-09-878-574-1462 (1-390)

QY 6 GluGluGluGlnValIleSerCysHisThrValGluGluTrpAsnAspGlnLeuGln 25  
DB 69 GAAAGAGGACAGTCATCGCGCTCCACACCGTTGATGAGTGAGACTGCACATCCAGAAAT 128  
QY 26 GlyAsnGluSerLysLysLeuValValAspPheThrAlaSerTrpCysGlyProCys 45  
DB 129 GCAAAAGACTCCAAAACATGATTTGTTGATTTTACTGCTTCCTGCTGATGCATGC 188  
QY 46 ArgPheIleAlaProPheLeuAlaGluLeuAlaLysLysPheThrSerValIlePheLeu 65  
DB 189 CGTTTATGCCCCCAGTTCCTTGACAGATTCAGAAAGAACTCCTGATGATCTTCCTC 248  
QY 66 LysValAspValAspGluLeuLysSerValSerGlnAspTrpAlaIleGluAlaMetPro 85  
DB 249 AAGTCGATGTGATGATGAGAGGCTGTTGCTGAGAGAAATATTCATTTAGGCGCATGCCA 308  
QY 86 ThrPheValPheValLysGluGlyThrLeuLeuAspLysValValGlyAlaLysAsp 105  
DB 309 ACCTTCCTCTTCTTGAAGAGATGCGAGATGCTGCAGCAAGGTGTTGCTACTAGATGAT 368  
QY 106 GluLeuGlnLysIle 111

```
/ SOFTWARE: Patentin Ver. 2.1
/
/   ID NO 5
/   LENGTH: 3129
/   TYPE: DNA
/   ORGANISM: Arabidopsis thaliana
/   FEATURE:
/   NAME/KEY: CDS
/   LOCATION: (1555)..(1896)
/   US-09-897-898-5

Alignment Scores:
Pred. No.: 1,06e-60      Length: 3129
Score: 482.00           Matches: 88
Percent Similarity: 89.38% Conservative: 13
Best Local Similarity: 77.88% Mismatches: 12
Query Match: 78.378      Indels: 0
DB: 9                  Gaps: 0

US-09-786-715-6 (1-120) x US-09-897-898-5 (1-3129)
QY 4 SerSerGluGluGlyGlnValIleSerCysHisThValGluGluTrpAsnAspGlnLeu 23
    ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
DB 1558 GCTTCGGAAGAGGACAAAGTATCGCTCCACACCGTTGAGACATGAGACGACGCTT 1617
QY 24 GlnLysGlyAsnGluSerLysLysLeuIleValAspPheThrAlaSerTrpCysGly 43
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 1618 CAGAAAGCTAATGATCCAAACTCTTGCTGCTGATTTACGCGCTCTTGCTGGA 1677
QY 44 ProCysArgPheIleAlaProPheLeuAlaGluLeuAlaLysPheThrSerValIle 63
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 1678 CATATGCGTTTCATCGCTCCATCTTCTTGCTGATTTGCTAGCAAACTTCTTACGCTT 1737
QY 64 PheLeuLysValAspValAspGluLeuLysSerValSerGlnAspTrpAlaIleGluAla 83
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 1738 TTCCTCAAGCTTGATCTGATGATTAATGAAGTCGGTGCAGAGTATGGCGATACAGGCG 1797
QY 84 MetProThrPheValPheValLysGluGlyThrLeuLeuAspLysValAlaGlyAlaLys 103
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 1798 ATGCCAACCTTCATGTTTGAAGGAGGAGATTTGGACAAAGTTGTGGAGCCAG 1857
QY 104 LysAspGluLeuGlnGlnLysIleGlnLysHisValAla 116
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 1858 AAAGATGAGCTTCAGTCACTACATTCGCAAAACACTTGCGT 1896

RESULT 4
/ US-09-897-898-7
/ Sequence 7, Application US/09897898
/ Patent No. US20020037303A1
/ GENERAL INFORMATION:
/ APPLICANT: DECKERS, HARM M.
/ APPLICANT: VAN ROOIJEN, GJOS
/ APPLICANT: BOOTHIE, JOSEPH
/ APPLICANT: GOLU, JANIS
/ APPLICANT: MOLONEY, MAURICE M.
/ APPLICANT: DALMIA, BIPIN K.
/ TITLE OF INVENTION: THIOREDOXIN AND THIOREDOXIN REDUCTASE CONTAINING OIL
/ FILE REFERENCE: 034547/0104
/ CURRENT APPLICATION NUMBER: US/09/897,898
/ CURRENT FILING DATE: 2001-09-21
/ PRIOR APPLICATION NUMBER: 09/577,147
/ PRIOR FILING DATE: 2000-05-24
/ PRIOR APPLICATION NUMBER: 09/448,600
/ PRIOR FILING DATE: 1999-11-24
/ PRIOR APPLICATION NUMBER: 09/084,777
/ PRIOR FILING DATE: 1998-05-27
/ PRIOR APPLICATION NUMBER: 60/047,753
/ PRIOR FILING DATE: 1997-05-27
/ PRIOR APPLICATION NUMBER: 60/047,779
/ PRIOR FILING DATE: 1997-05-28
/ PRIOR APPLICATION NUMBER: 60/075,863
/ PRIOR FILING DATE: 1998-02-25
/ PRIOR APPLICATION NUMBER: 60/075,864
```

```
/ PRIOR FILING DATE: 1998-02-25
/ NUMBER OF SEQ ID NOS: 24
/ SOFTWARE: Patentin Ver. 2.1
/
/   ID NO 7
/   LENGTH: 3888
/   TYPE: DNA
/   ORGANISM: Unknown Organism
/   FEATURE:
/   NAME/KEY: CDS
/   LOCATION: (1555)..(1908)
/   NAME/KEY: CDS
/   LOCATION: (2149)..(2655)
/   OTHER INFORMATION: Description of Unknown Organism: Phaseolin
/   US-09-897-898-7

Alignment Scores:
Pred. No.: 1.47e-60      Length: 3888
Score: 482.00           Matches: 88
Percent Similarity: 89.38% Conservative: 13
Best Local Similarity: 77.88% Mismatches: 12
Query Match: 78.378      Indels: 0
DB: 9                  Gaps: 0

US-09-786-715-6 (1-120) x US-09-897-898-7 (1-3888)
QY 4 SerSerGluGluGlyGlnValIleSerCysHisThValGluGluTrpAsnAspGlnLeu 23
    ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
DB 2317 GCTTCGGAAGAGGACAAAGTATCGCTCCACACCGTTGAGACATGAGACGACGCTT 2376
QY 24 GlnLysGlyAsnGluSerLysLysLeuIleValAspPheThrAlaSerTrpCysGly 43
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 2377 CAGAAAGCTAATGATCCAAACTCTTGCTGCTGATTTACGCGCTCTTGCTGGA 2436
QY 44 ProCysArgPheIleAlaProPheLeuAlaGluLeuAlaLysPheThrSerValIle 63
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 2437 CATGTCGTTTCATCGCTCCATCTTCTTGCTGATTTGCTAGCAAACTTCTTACGCTT 2496
QY 64 PheLeuLysValAspValAspGluLeuLysSerValSerGlnAspTrpAlaIleGluAla 83
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 2497 TTCCTCAAGCTTGATCTGATGATTAATGAAGTCGGTGCAGAGTATGGCGATACAGGCG 2556
QY 84 MetProThrPheValPheValLysGluGlyThrLeuLeuAspLysValAlaGlyAlaLys 103
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 2557 ATGCCAACCTTCATGTTTGAAGGAGGAGATTTGGACAAAGTTGTGGAGCCAG 2616
QY 104 LysAspGluLeuGlnGlnLysIleGlnLysHisValAla 116
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 2617 AAAGATGAGCTTCAGTCACTACATTCGCAAAACACTTGCGT 2655

RESULT 5
/ US-09-878-574-3723
/ Sequence 3723, Application US/09878574
/ Patent No. US20020110548A1
/ GENERAL INFORMATION:
/ APPLICANT: BYNUM, Joseph R.
/ APPLICANT: LA ROSA, Thomas J.
/ APPLICANT: THOMPSON, Michael D.
/ TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated with
/ FILE REFERENCE: 38-21(15401)B
/ CURRENT APPLICATION NUMBER: US/09/878,574
/ CURRENT FILING DATE: 2001-12-21
/ PRIOR APPLICATION NUMBER: 09/333,535
/ PRIOR FILING DATE: 1999-06-14
/ NUMBER OF SEQ ID NOS: 15775
/ SEQ ID NO 3723
/ LENGTH: 402
/ TYPE: DNA
/ ORGANISM: Glycine max
/ OTHER INFORMATION: Clone ID: LIB3028-006-Q1-B1-F8
/ US-09-878-574-3723
```

;; PRIOR FILING DATE: 1997-05-28  
;; PRIOR APPLICATION NUMBER: 60/075,863  
;; PRIOR FILING DATE: 1998-02-25  
;; PRIOR APPLICATION NUMBER: 60/075,864  
;; PRIOR FILING DATE: 1998-02-25  
;; NUMBER OF SEQ ID NOS: 24  
;; SOFTWARE: PatentIn Ver. 2.1  
;; SEQ ID NO 10  
;; LENGTH: 3888  
;; TYPE: DNA  
;; ORGANISM: Unknown Organism  
;; FEATURE:  
;; NAME/KEY: CDS  
;; LOCATION: (1555)..(2250)  
;; NAME/KEY: CDS  
;; LOCATION: (2491)..(2655)  
;; OTHER INFORMATION: Description of Unknown Organism: Phaseolin  
;; OTHER INFORMATION: promoter-Trrh oleosin-phaseolin terminator  
US-09-897-898-10

Alignment Scores:  
Pred. No.: 5,26e-61 Length: 3888  
Score: 485.00 Matches: 89  
Percent Similarity: 88.70% Conservative: 13  
Best Local Similarity: 77.39% Mismatches: 13  
Query Match: 78.86% Indels: 0  
Gaps: 0

US-09-786-715-6 (1-120) x US-09-897-898-10 (1-3888)

QY 4 SerSerGlUGlUGlGlnValIIleSerCysHisThrValGluGluTrpPasnAspGlnLeu 23  
Db 1558 GCTTCGGAAGAGCAAGATGATGCGCTGCCACACCTTGAGACATGGAACGACACCTT 1617  
QY 24 GlnTsgLysnGlnSerLysLysLeuIIeValValAspPheThrAlaSerTrpCysGly 43  
Db 1618 CAGAGGCTATGATGATCAAAACCTTGTTGCTGATTTTCACGGCTTCTTGCTGATGGA 1677  
QY 44 ProcysArpPheIIeAlaProPheLeuAlaGluLeuAlaLysLysPheThrSerValIle 63  
Db 1678 CCATGTCGTTTCATCGCTCCATCTTCTTGCTGATTTGGCTAAGAACTTCTTAACGTCCT 1737  
QY 64 PheLeuLysValAspValAspGluLeuLysSerValSerGlnAspTrpAlaIleGluAla 83  
Db 1738 TTCTCAAGGTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1797  
QY 84 MetProthrPheValPheValLysGluGlyThreLeuLysPheValValAlaLys 103  
Db 1798 ATGCCAACCTTCATGTTTGAAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1857  
QY 104 LysAspGluLeuGlnGlnLysIleGlnLysHisValAlaSerAla 118  
Db 1858 AAGATGAGCTTCAGTCTACCATTTGCCAACACCTTGCTATGGCG 1902

RESULT 2  
US-09-938-842A-163  
; Sequence 163, Application US/09938842A  
; Patent No. US20020160378A1  
; GENERAL INFORMATION:  
; APPLICANT: Harper, Jeff  
; APPLICANT: Kieps, Joel  
; APPLICANT: Wang, Xun  
; APPLICANT: Zhu, Tong  
; TITLE OF INVENTION: STRESS-REGULATED GENES OF PLANTS, TRANSGENIC PLANTS CONTAINING  
; FILE REFERENCE: SCRIP1300-3  
; CURRENT APPLICATION NUMBER: US/09/938,842A  
; PRIOR FILING DATE: 2001-08-24  
; PRIOR APPLICATION NUMBER: US 60/227,866  
; PRIOR FILING DATE: 2000-08-24  
; PRIOR APPLICATION NUMBER: US 60/264,647  
; PRIOR FILING DATE: 2001-01-16  
; PRIOR APPLICATION NUMBER: US 60/300,111

;; PRIOR FILING DATE: 2001-06-22  
;; NUMBER OF SEQ ID NOS: 5379  
;; SEQ ID NO 163  
;; LENGTH: 345  
;; TYPE: DNA  
;; ORGANISM: Arabidopsis thaliana  
US-09-938-842A-163

Alignment Scores:  
Pred. No.: 4,05e-62 Length: 345  
Score: 482.00 Matches: 88  
Percent Similarity: 89.38% Conservative: 13  
Best Local Similarity: 77.88% Mismatches: 12  
Query Match: 78.37% Indels: 0  
Gaps: 0

US-09-786-715-6 (1-120) x US-09-938-842A-163 (1-345)

QY 4 SerSerGlUGlUGlGlnValIIleSerCysHisThrValGluGluTrpPasnAspGlnLeu 23  
Db 4 GCTTCGGAAGAGCAAGATGATGCGCTGCCACACCTTGAGACATGGAACGACACCTT 63  
QY 24 GlnTsgLysnGlnSerLysLysLeuIIeValValAspPheThrAlaSerTrpCysGly 43  
Db 64 CAGAGGCTATGATGATCAAAACCTTGTTGCTGATTTTCACGGCTTCTTGCTGATGGA 123  
QY 44 ProcysArpPheIIeAlaProPheLeuAlaGluLeuAlaLysLysPheThrSerValIle 63  
Db 124 CCATGTCGTTTCATCGCTCCATCTTCTTGCTGATTTGGCTAAGAACTTCTTAACGTCCT 183  
QY 64 PheLeuLysValAspValAspGluLeuLysSerValSerGlnAspTrpAlaIleGluAla 83  
Db 184 TTCTCAAGGTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 243  
QY 84 MetProthrPheValPheValLysGluGlyThreLeuLysPheValValAlaLys 103  
Db 244 ATGCCAACCTTCATGTTTGAAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 303  
QY 104 LysAspGluLeuGlnGlnLysIleGlnLysHisValAla 116  
Db 304 AAGATGAGCTTCAGTCTACCATTTGCCAACACCTTGCT 342

RESULT 3  
US-09-897-898-5  
; Sequence 5, Application US/09897898  
; Patent No. US20020037303A1  
; GENERAL INFORMATION:  
; APPLICANT: DECKERS, HARM M.  
; APPLICANT: VAN ROOIJEN, GIJS  
; APPLICANT: BOOTHE, JOSEPH  
; APPLICANT: GOUL, JANIS  
; APPLICANT: MOLONEY, MAURICE M.  
; APPLICANT: PALMIA, BIPIN K.  
; TITLE OF INVENTION: THIOREDOXIN AND THIOREDOXIN REDUCTASE CONTAINING OIL  
; FILE REFERENCE: 034547/0104  
; CURRENT APPLICATION NUMBER: US/09/897,898  
; CURRENT FILING DATE: 2001-09-21  
; PRIOR APPLICATION NUMBER: 09/577,147  
; PRIOR FILING DATE: 2000-05-24  
; PRIOR APPLICATION NUMBER: 09/448,600  
; PRIOR FILING DATE: 1999-11-24  
; PRIOR APPLICATION NUMBER: 09/084,777  
; PRIOR FILING DATE: 1998-05-27  
; PRIOR APPLICATION NUMBER: 60/047,753  
; PRIOR FILING DATE: 1997-05-27  
; PRIOR APPLICATION NUMBER: 60/047,779  
; PRIOR FILING DATE: 1997-05-28  
; PRIOR APPLICATION NUMBER: 60/075,863  
; PRIOR FILING DATE: 1998-02-25  
; PRIOR APPLICATION NUMBER: 60/075,864  
; PRIOR FILING DATE: 1998-02-25  
; NUMBER OF SEQ ID NOS: 24

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - nucleic search, using frame\_plus\_p2n model

Run on: August 17, 2003, 19:25:47 ; Search time 139.876 seconds  
(without alignments)  
1919.938 Million cell updates/sec

Title: US-09-786-715-6  
Perfect score: 615  
Sequence: 1 MAGSSEGVISCHTEWEMN.....GAKKDELQKIQKHVASASA 120

Scoring table: BLOSUM62  
Xgapop 10.0, Xgapext 0.5  
Ygapop 10.0, Ygapext 0.5  
Fgapop 6.0, Fgapext 7.0  
Delop 6.0, Delext 7.0

Searched: 1504479 seqs, 1118970152 residues  
Total number of hits satisfying chosen parameters: 3008958

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000  
Post-processing: Minimum Match 08  
Maximum Match 100%  
Listing first 45 summaries

Command line parameters:  
-MODEL=frame+p2n.model -DEV=xlp  
-Q=/cgn2\_1/USPTO\_SPOOL/US09786715/runat\_11082003\_150515\_6129/app\_query.fasta\_1.1052  
-DB=Published\_Applications\_NA -QMT=fastap -SUFFIX=trpb -MINMATCH=0.1  
-LOOPEL=0 -LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blosum62  
-FRAN=human40.cdi -LIST=45 -DOCALLIGN=200 -THR\_SCORE=pcr -THR\_MAX=100  
-THR\_MIN=0 -ALIGN=15 -MODE=local -OUTFMT=ptc -NORM=ext -HEARSIZE=500 -MINLEN=0  
-MAXLEN=2000000000 -USBR=US0987615 @cgn\_1.1.560 @runat\_11082003\_150515\_6129  
-NCPU=6 -ICPU=3 -NO\_MMAP -LARGEOUTPUT -NEG\_SCORES=0 -WAIT -DSFBLOCK=100  
-LONGLOG -DEV\_TIMEOUT=120 -WARN\_TIMEOUT=30 -THREDS=1 -XGAPOP=10 -XGAPEXT=0.5  
-FCAPOP=6 -FCAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :  
1: /cgn2\_6/ptodata/2/pubpna/US07\_PUBCOMB.seq:\*  
2: /cgn2\_6/ptodata/2/pubpna/PCT\_NEW\_PUB.seq:\*  
3: /cgn2\_6/ptodata/2/pubpna/US06\_NEW\_PUB.seq:\*  
4: /cgn2\_6/ptodata/2/pubpna/US06\_PUBCOMB.seq:\*  
5: /cgn2\_6/ptodata/2/pubpna/US07\_NEW\_PUB.seq:\*  
6: /cgn2\_6/ptodata/2/pubpna/PCTUS\_PUBCOMB.seq:\*  
7: /cgn2\_6/ptodata/2/pubpna/US08\_NEW\_PUB.seq:\*  
8: /cgn2\_6/ptodata/2/pubpna/US08\_PUBCOMB.seq:\*  
9: /cgn2\_6/ptodata/2/pubpna/US09\_PUBCOMB.seq:\*  
10: /cgn2\_6/ptodata/2/pubpna/US09B\_PUBCOMB.seq:\*  
11: /cgn2\_6/ptodata/2/pubpna/US09C\_PUBCOMB.seq:\*  
12: /cgn2\_6/ptodata/2/pubpna/US09\_NEW\_PUB.seq:\*  
13: /cgn2\_6/ptodata/2/pubpna/US10A\_PUBCOMB.seq:\*  
14: /cgn2\_6/ptodata/2/pubpna/US10B\_PUBCOMB.seq:\*  
15: /cgn2\_6/ptodata/2/pubpna/US10\_NEW\_PUB.seq:\*  
16: /cgn2\_6/ptodata/2/pubpna/US60\_NEW\_PUB.seq:\*  
17: /cgn2\_6/ptodata/2/pubpna/US60\_PUBCOMB.seq:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	485	78.9	3888	9	US-09-897-898-10 Sequence 10, Appl

2	482	78.4	345	10	US-09-938-842A-163	Sequence 163, App
3	482	78.4	3129	9	US-09-897-898-5	Sequence 5, Appl1
4	482	78.4	3888	9	US-09-897-898-7	Sequence 7, Appl1
5	438	71.2	402	10	US-09-878-574-3723	Sequence 3723, Ap
6	424	68.9	390	10	US-09-878-574-104	Sequence 104, App
7	391	63.6	390	10	US-09-878-574-1662	Sequence 1462, Ap
8	387	62.9	540	12	US-10-349-782-12	Sequence 12, Appl
9	381.5	62.0	360	10	US-09-938-842A-164	Sequence 164, Appl
10	376.5	61.2	357	10	US-09-938-842A-1736	Sequence 1736, Ap
11	376.5	61.2	357	14	US-10-323-362-1	Sequence 1, Appl1
12	368	59.8	393	12	US-10-091-841-5	Sequence 5, Appl1
13	367	59.7	382	12	US-10-091-841-3	Sequence 3, Appl1
14	361	58.7	369	12	US-10-194-885-8	Sequence 8, Appl1
15	361	58.7	369	12	US-10-091-841-1	Sequence 1, Appl1
16	358	58.2	659	12	US-10-349-782-9	Sequence 9, Appl1
17	316	51.4	277	10	US-09-878-574-8135	Sequence 8135, Ap
18	300	48.8	260	10	US-09-878-574-12941	Sequence 12941, A
19	293	47.6	370	10	US-09-878-574-3169	Sequence 3169, Ap
20	289	47.0	268	10	US-09-878-574-5995	Sequence 5995, Ap
21	285.5	46.4	402	10	US-09-938-842A-225	Sequence 225, App
22	285.5	46.4	613	9	US-09-770-149-855	Sequence 855, App
23	283	46.0	270	10	US-09-878-574-13445	Sequence 13445, A
24	279	45.4	272	10	US-09-878-574-12181	Sequence 11, Appl
25	274	44.6	297	12	US-10-349-782-11	Sequence 677, App
26	272	44.2	456	10	US-09-924-035A-677	Sequence 1288, Ap
27	256	41.6	350	9	US-09-938-842A-2688	Sequence 2688, Ap
28	255	41.5	390	10	US-09-878-574-13137	Sequence 13137, A
29	252	41.0	267	10	US-09-878-574-1137	Sequence 7974, Ap
30	237	38.5	261	10	US-09-878-574-7974	Sequence 914, App
31	229	37.2	440	9	US-09-770-444-914	Sequence 147, App
32	226	36.7	357	10	US-09-938-842A-147	Sequence 76, Appl
33	222	36.1	301	9	US-09-815-343-76	Sequence 1366, Ap
34	222	36.1	413	14	US-10-102-524-1396	Sequence 3641, Ap
35	222	36.1	418	10	US-09-796-692-3641	Sequence 3641, Ap
36	222	36.1	418	14	US-10-040-862-3641	Sequence 1059, Ap
37	222	36.1	464	10	US-09-998-598-1059	Sequence 1145, Ap
38	222	36.1	479	11	US-09-535-459-1145	Sequence 65, Appl
39	222	36.1	536	14	US-10-066-543-65	Sequence 1247, Ap
40	222	36.1	540	10	US-09-920-300A-1247	Sequence 1247, Ap
41	222	36.1	540	13	US-10-033-528-1247	Sequence 88, Appl
42	222	36.1	557	10	US-09-884-441-88	Sequence 88, Appl
43	222	36.1	557	11	US-09-907-965-88	Sequence 88, Appl
44	222	36.1	557	14	US-10-198-053-88	Sequence 1139, Ap
45	222	36.1	568	11	US-09-535-459-1139	

## ALIGNMENTS

RESULT 1  
US-09-897-898-10  
; Sequence 10, Application US/09897898  
; Patent No. US20020037303A1  
; GENERAL INFORMATION:  
; APPLICANT: DECKERS, HARM M.  
; APPLICANT: VAN ROOIJEN, GIJS  
; APPLICANT: BOOTHIE, JOSEPH  
; APPLICANT: GOLL, JANIS  
; APPLICANT: MOLONEY, MAURICE M.  
; APPLICANT: DALMIA, BIPIN K.  
; TITLE OF INVENTION: THIOREDOXIN AND THIOREDOXIN REDUCTASE CONTAINING OIL  
; FILE REFERENCE: 034547/0104  
; CURRENT APPLICATION NUMBER: US/09/897,898  
; CURRENT FILING DATE: 2001-09-21  
; PRIOR APPLICATION NUMBER: 09/577,147  
; PRIOR FILING DATE: 2000-05-24  
; PRIOR APPLICATION NUMBER: 09/448,600  
; PRIOR FILING DATE: 1999-11-24  
; PRIOR APPLICATION NUMBER: 09/084,777  
; PRIOR FILING DATE: 1998-05-27  
; PRIOR APPLICATION NUMBER: 60/047,753  
; PRIOR FILING DATE: 1997-05-27  
; PRIOR APPLICATION NUMBER: 60/047,779



GenCore version 5.1.6  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - nucleic search, using frame\_plus\_p2n model

Run on: August 17, 2003, 18:47:50 ; Search time 1887.34 Seconds

(without alignments)  
2644.453 Million cell updates/sec

Title: US-09-786-715-2

Perfect score: 637

Sequence: 1 MAEEGYIACHKIDWEGQL.....GLSKILELSTTPATSTA 122

Scoring table:

BLOSUM62	
Xgapop 10.0 , Xgapext 0.5	
Ygapop 10.0 , Ygapext 0.5	
Fgapop 6.0 , Fgapext 7.0	
Delop 6.0 , Delext 7.0	

Searched: 2888711 seqs, 2045481386 residues

Total number of hits satisfying chosen parameters: 5777422

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame\_plus\_p2n.model -DEV=slp  
-O=/cgn2.1/USPTO/spool/US09786715/runat\_11082003\_150514\_6035/app\_query.fasta\_1.1052  
-DB=GenEmbl -QFMT=fastap -SUFFIX=rge -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0  
-UNITS=bits -START=1 -END=1 -MATRIX=blsum62 -TRANS=humana40.cdi -LIST=45  
-DOCLIGN=200 -THR.SCORE=pct -THR.MAX=100 -THR.MIN=0 -ALIGN=15 -MODE=LOCAL  
-OUTFMT=pic -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000  
-USER=us09786715 -GCGN\_1.1\_5721 -ernat\_11082003\_150514\_6035 -NCPU=6 -ICPU=3  
-NO\_MMAP -LARGOQUERY -NEG.SCORES=0 -WAIT -DSPBLOCK=100 -XGAPOP=6 -FGAPOP=6  
-DEV.TIMEOUT=120 -WARN.TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6  
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

1:	GenEmbl:*
2:	gb_da:*
3:	gb_hcg:*
4:	gb_in:*
5:	gb_om:*
6:	gb_ov:*
7:	gb_ph:*
8:	gb_pl:*
9:	gb_pr:*
10:	gb_ro:*
11:	gb_sts:*
12:	gb_sy:*
13:	gb_un:*
14:	gb_vi:*
15:	em_da:*
16:	em_fun:*
17:	em_hum:*
18:	em_in:*
19:	em_mus:*
20:	em_or:*
21:	em_om:*
22:	em_ov:*
23:	em_pat:*
24:	em_ph:*
25:	em_pl:*
26:	em_ro:*
27:	em_sts:*
28:	em_un:*

29: em\_vl:\*

30: em\_htg\_hum:\*

31: em\_htg\_inv:\*

32: em\_htg\_other:\*

33: em\_htg\_mus:\*

34: em\_htg\_pln:\*

35: em\_htg\_rnd:\*

36: em\_htg\_nam:\*

37: em\_htg\_vrt:\*

38: em\_sy:\*

39: em\_htgo\_hum:\*

40: em\_htgo\_mus:\*

41: em\_htgo\_other:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	385	60.4	345	8	AF483265 Populus t
2	377	59.2	664	8	AF323593 Prunus pe
3	376	59.0	345	6	AX505468 Sequence
4	376	59.0	480	8	AY088687 Arabidops
5	376	59.0	497	8	ATTHIOARA
6	368	57.8	603	8	RCFHIORXN
7	354	55.6	665	8	AY170650
8	352	55.3	595	8	D87984
9	352	55.3	630	8	PSA310990
10	348.5	54.7	698	8	NTTRNA
11	348.5	54.4	590	8	AY088698
12	345.5	54.2	360	6	AX505469
13	345.5	54.2	360	6	BT004710
14	345.5	54.2	556	8	AK118542
15	343.5	53.9	641	8	AP051206
16	343	53.8	526	11	G73679
17	343	53.8	601	8	OS092541
18	343	53.8	686	6	E08194
19	343	53.8	687	8	RICHT
20	342.5	53.8	642	8	ATTHIREDD1
21	342.5	53.8	784	8	AY271308
22	337	52.9	357	8	AY040028
23	337	52.9	528	8	ATTHIREDD4
24	337	52.9	530	8	AK118035
25	337	52.9	560	8	AY087159
26	337	52.9	561	8	AF360227
27	335.5	52.7	653	6	AR016869
28	335.5	52.7	653	6	AR020895
29	335.5	52.7	653	6	AR027218
30	335.5	52.7	653	6	AR038505
31	335.5	52.7	653	6	AR064647
32	335.5	52.7	653	6	AR067572
33	335.5	52.7	653	6	I38524
34	335.5	52.7	653	6	I56999
35	335.5	52.7	653	6	I59865
36	335.5	52.7	653	6	I75192
37	334	52.4	593	8	PSA319808
38	334	51.8	582	3	AB061204
39	323	50.7	357	6	AX412276
40	323	50.7	357	6	AX412277
41	323	50.7	357	6	AX412278
42	323	50.7	357	6	AX412279
43	323	50.7	357	6	AX412473
44	323	50.7	357	6	AX412474
45	323	50.7	357	6	AX412566

## ALIGNMENTS

RESULT 1

LOCUS	AF483265	345 bp	mRNA	linear	PLN 01-APR-2002
DEFINITION	Populus tremula x Populus tremuloides thioresdoxin H mRNA, complete cds.				
ACCESSION	AF483265				
VERSION	AF483265.1				
KEYWORDS	GI:19851971				
SOURCE	Populus tremula x Populus tremuloides				
ORGANISM	Populus tremula x Populus tremuloides Eukaryota; Viridiplantae; Streptophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids; eurosids I; Malpighiales; Salicaceae; Populus.				
REFERENCE	1 (bases 1 to 345) Behm, M. and Jacquot, J.P.				
AUTHORS	Isolation and characterization of thioresdoxin h from poplar xylem				
JOURNAL	Plant Physiol. Biochem. 38 (5), 363-369 (2000)				
TITLE	2 (bases 1 to 345) Behm, M. and Jacquot, J.P.				
REFERENCE	Direct Submission				
JOURNAL	Submitted (13-Feb-2002) Faculte des Sciences, Universite Henri Poincare, Bd des Aiguillettes, Vandoeuvre 54506, France				
FEATURES	Location/Qualifiers				
SOURCE	1..345				
CDS	/organism="Populus tremula x Populus tremuloides"				
	/mol_type="mRNA"				
	/db_xref="taxon:47664"				
	1..345				
	/codon_start=1				
	/product="thioresdoxin H"				
	/protein_id="AA199941.1"				
	/db_xref="GI:19851972"				
	/translation="MAERGQVYIACHTVDTWKHEKSGSKLIVDFTSWCPCKMKIAPFPAELAKKFPNVTFLKVDVDELKVAEEMNVEAMPTFILDKGLVDKTVGADKDKGLPILVAKHATA"				
BASE COUNT	101 a	61 c	95 g	88 t	
ORIGIN					
Alignment Scores:					
Pred. No.:	1.22e-35	Length:	345		
Score:	385.00	Matches:	70		
Percent Similarity:	78.85%	Conservative:	12		
Best local Similarity:	67.31%	Mismatches:	22		
Query Match:	60.44%	Indels:	0		
DB:	8	Gaps:	0		
US-09-786-715-2 (1-122) x AF483265 (1-345)					
OY	1 Melalaglugluglglhvalilialacyshtslystleaapglutrgluglglneu 20				
DB	1 ATGGCCGAGAGAGACAAAGTTATGTCCTGCCACACAGTGGATACCTGGAAAGACATTTTC 60				
OY	21 Glylystrplyaspserserglulsylneuvalvalasphehrhalaasertrpcysgly 40				
DB	61 GAGAGAGGAAAGGCTCTCAGAAACTGATGTGTGCTGATTTACTGCTTCATGGTGTCCA 120				
DB	41 Procyasargalallealeaertrpphenhrcjluenualysasanaenproasvalala 60				
OY	121 CCATGTAAATGATGTCCTCCAACTCTGCCACAGTTGGGAGAAAGTTCCCAATGTACA 180				
DB	61 PheleuysValaspvalaspglulneuanserValalaserLysTrpGluIleasAla 80				
DB	181 TTCTTGAAGGTGAGTGGATGATGAMTTGAAGGCTGTCTGTAGAGAGTGAATGTGAGGCA 240				
OY	81 MetcProthrPheValPheleuLysGlyLysIleIleGluLysIleValGlyAlaasp 100				
DB	241 ATGCCCACTTTATTTCTCTGAAAGATGAAATTAAGTGCACAAACTGTGCTGTGAT 300				
OY	101 LysValGlyLeu 104				
DB	301 AAAGATGGCCTG 312				
RESULT 2					
AF323593					

LOCUS	AF323593	664 bp	mRNA	linear	PLN 02-NOV-2003
DEFINITION	Prunus persica thiorodoxin H mRNA, complete cds.				
ACCESSION	AF323593				
VERSION	AF323593.1	GI:16588842			
KEYWORDS					
SOURCE	Prunus persica (peach)				
ORGANISM	Prunus persica				
REFERENCE	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids; eurosids I; Rosales; Rosaceae; Amygdaloideae; Prunus.				
AUTHORS	Callahan,A.M., Morgens,P.H. and Cohen,R.A.				
TITLE	Isolation and initial characterization of cDNAs for mRNAs regulated during peach fruit development				
JOURNAL	J. Am. Soc. Hortic. Sci. 118, 531-537 (1993)				
AUTHORS	2. (bases 1 to 664)				
TITLE	Callahan,A.M., Morgens,P.H., Cohen,R.A. and Scorza,R.				
REFERENCE	Regulation of peach gene expression in a peach/almond hybrid				
AUTHORS	Unpublished				
TITLE	3. (bases 1 to 664)				
JOURNAL	Callahan,A.M., Dunn,L.L. and Cohen,R.A.				
FEATURES	Direct Submission				
SOURCE	Submitted (22-NOV-2000) USDA-ARS, Appalachian Fruit Research Station, 45 Wiltshire Rd., Kearneysville, WV 25430, USA				
location/Qualifiers	1..664				
/organism="Prunus persica"					
/mol_type="mRNA"					
/cultivar="B612615"					
/db_xref="taxon:3760"					
/clone="pch306c"					
/tissue_type="ripe fruit"					
71..481					
/codon_start=1					
/product="thiorodoxin H"					
/protein_id="AA28915.1"					
/db_xref="GI:16588843"					
/translation="MAEENQYIGCHTQAMDEQLHKGNENKLLVYVDFATMCGPCRIIAPLIELAKKTPVETFLKIVVDLRIVSEMGVEMPTFLFKEGKIVDYVYVAKKKKELQIKVAKHVAANAASATASATATATATASA"					
BASE COUNT	182 a	147 c	161 g	174 t	
ORIGIN					
Alignment Scores:					
Pred. NO.:	2.19e-34	Length:	664		
Score:	377.00	Matches:	74		
Percent Similarity:	68.18%	Conservative:	16		
Best Local Similarity:	56.06%	Mismatches:	32		
Query Match:	59.18%	Indels:	10		
DB:	8	Gaps:	1		
US-09-786-715-2 (1-122) x AF323593 (1-664)					
OY	1	MetaIaGluGluGlyGlnValIItIleAlaCySHisLysIleAspGluTrpGluGlyGlnLeu	20		
Db	71	ATGGCGGAGGAAATCAAGTATCGGCTGCCACACTACTCAAGCTGGGAAGACAGCTC	130		
OY	21	GlyLysTrpLysAspSerGlnLysLeuValValAspPheThrAlaSerTrpCysGly	40		
Db	131	CATAAGGGAACGAGACAAGAAATCGTGTGGTGTGATTCACGGCTTCGTGTGTGA	190		
OY	41	ProCysArgAlaIleAlaIleProTyrPheThrGlnLeuAluLysAsnAsnProAsnValAla	60		
Db	191	CCGTGCGCGTGTGATCGCCCAATCTGTGGCGGAGTTGGCTAAAGAGACCCCAAGATGACG	250		
OY	61	PheLeuLysValaAspValaAspGluLeuAsnSerValaIaSerLysTrpGluIleAsnAla	80		
Db	251	TTCTTAAGGTGAGCGAGTGAAGTACAGTCTTTTCCGAGAGATGGGCGTGTGAGGCA	310		
OY	81	MetProThrPheValPheLeuLysGlyLysIleIleGluLysIleValAlaAsp	100		
Db	311	ATGGCTACCTTCCTCTCTCAAGAGAGGCAAGATAGTGGACAAGGTTGTGGTGCCAAAG	370		

```

Oy 101 LysValGlyLeuSerLysLysIleLeuGlu----- 110
Db 371 AAGACGAGTACAGATCAAGATGCGCGCTGCCGCCGCTCGCTCCGCC 430
Oy 111 LeuSerGlyThrProAlaIaIaThSerThrala 122
Db 431 ACCTCTGCTCTGCGACCGCTGTACACTACTGCT 466

RESULT 3
AX505468 345 bp DNA linear PAT 27-SEP-2002
DEFINITION Sequence 163 from Patent WO0216655.
ACCESSION AX505468
VERSION AX505468.1
KEYWORDS GI:23386705
SOURCE Arabidopsis thaliana (thale cress)
ORGANISM Arabidopsis thaliana
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsi.

REFERENCE
1 Harper, J.F., Kreps, J., Wang, X. and Zhu, T.
Stress-regulated genes of plants, transgenic plants containing
same, and methods of use
Patent: WO 0216655-A 163, 28-FEB-2002;
The Scripps Research Institute (US) ; Syngenta Participations AG
(CH)

FEATURES
Source
1. 345
/organism="Arabidopsis thaliana"
/mol_type="genomic DNA"
/db_xref="taxon:3702"

BASE COUNT 90 a 68 c 90 g 97 t

ORIGIN
Alignment Scores:
Pred. No.: 1,34e-34 Length: 345
Score: 376.00 Matches: 71
Percent Similarity: 77.57% Conservative: 12
Best Local Similarity: 66.36% Mismatches: 24
Query Match: 59.03% Indels: 0
Gaps: 0

US-09-786-715-2 (1-122) x AX505468 (1-345)

Oy 2 AlaGluGluGlyGlnValIleAlaCysHisLysIleAspGluTTPGluGlyGlnLeuGly 21
Db 7 TCGGAAGAGAGACAGATGATCGCTGCCACACCGTTCAGACATGGAGACGACGCTTCAG 66
Oy 22 LysTTPlysAspSerGluLysLeuValValAspPheThrIaIaSerLysTTPCysGlyPro 41
Db 67 AAGGCTAATGATCAAAACCTCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 126
Oy 42 CysArgAlaIleAlaProTyrPheThrGluLeuValAlaLysAsnAsnProAsnValAlaPhe 61
Db 127 TGTGTTTCATCGCTCCATCTCTTGTGCTGATTTGGCTAGAAACTTCCTAACGCTGTTTC 186
Oy 62 LeuLysValAspValAspGluLeuAsnSerValAlaSerLysTTPGluIleAsnAlaMet 81
Db 187 CTCAGGTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 246
Oy 82 ProThrPheValPheLeuLysGlyLysIleIleGluLysIleValGlyValAlaAspLys 101
Db 247 CCAACCTTCATGTTTTCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 306
Oy 102 ValGlyLeuSerLysLysIle 108
Db 307 GATGAGCTTCAGTCTACCATT 327

RESULT 4
AY088687 480 bp mRNA linear PLN 14-APR-2003
LOCUS AY088687
DEFINITION Arabidopsis thaliana clone 9189 mRNA, complete sequence.

```

```

ACCESSION AY088687
VERSION AY088687.1
KEYWORDS GI:21407461
SOURCE FIL CDNA.
ORGANISM Arabidopsis thaliana (thale cress)
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsi.

REFERENCE
1 Haas, B.J., Volfovsky, N., Town, C.D., Troupkhan, M., Alexandrov, N.,
Feldmann, K.A., Flavell, R.B., White, O. and Salzberg, S.L.
Full-length messenger RNA sequences greatly improve genome
annotation
Genome Biol. 3 (6), RESEARCH0029 (2002)

JOURNAL MEDLINE
22088475
12093376
2 (bases 1 to 480)
Brover, V., Troupkhan, M., Alexandrov, N., Lu, Y.-P., Flavell, R. and
Feldmann, K.
Full-length cDNA from Arabidopsis thaliana
Unpublished
3 (bases 1 to 480)
Brover, V., Troupkhan, M., Alexandrov, N., Lu, Y.-P., Flavell, R. and
Feldmann, K.
Direct Submission
Submitted (11-MAR-2002) Ceres, Inc, 3007 Malibu Canyon Road,
Malibu, CA 90265, USA

COMMENT
This clone sequence is one of 5,000 Ceres full-length cDNAs made
available to TIGR and Genbank. The following quality assessment of
this set was done by comparison with known proteins: two percent of
the clones are estimated to be 5'-truncated; less than one percent
are 3'-truncated; approximately two percent represent alternative
splice variants, including unspliced introns and spliced exons; one
percent may contain premature stop codons; five percent may have
frame shifts in a coding region. A sequence is considered to be
5'-truncated if it lacks the translation initiation start (ATG). A
sequence is considered to be 3'-truncated if it lacks the
C-terminal end of the encoded protein. Please note that these cDNA
sequences are derived from the WS or Laer ecotypes and therefore
may contain polymorphisms when compared to sequences from Col-0.
GenSet carried out the library production and sequencing of the
full-length clones. Ceres, Inc. carried out the clustering of the
5' sequences, selection of clones, and sequence assembly.

FEATURES
Source
1. 480
/organism="Arabidopsis thaliana"
/mol_type="mRNA"
/db_xref="taxon:3702"
/clone="9189"

CDS
13..357
/codon_start=1
/product="Chlorodoxin h"
/protein_id="AAM67008.1"
/db_xref="GI:21617958"
/translation="MASEGGVIVICHTVEIWNEDLOKANEKTLVVVDFTSMGCPGR
FIAPRPADLAKKLINVLFLKVDTELKSVADMDIQAAMPFTPLFKKGIIDKYVAKK
DELSTIAKHIA"

BASE COUNT 124 a 93 c 111 g 152 t

ORIGIN
Alignment Scores:
Pred. No.: 1,96e-34 Length: 480
Score: 376.00 Matches: 71
Percent Similarity: 77.57% Conservative: 12
Best Local Similarity: 66.36% Mismatches: 24
Query Match: 59.03% Indels: 0
Gaps: 0

US-09-786-715-2 (1-122) x AY088687 (1-480)

Oy 2 AlaGluGluGlyGlnValIleAlaCysHisLysIleAspGluTTPGluGlyGlnLeuGly 21
Db 19 TCGGAAGAGAGACAGATGATCGCTGCCACACCGTTCAGACATGGAGACGACGCTTCAG 78

```

OY	22	LysTrpLysAspSerGlnLysLeuValValValAspPheThrAlaSerTrpGlyPro	41
		::	
	79	AAGCGTATGAAATCCAAACACTCTTGSGTGTTACACGCTTTGTTGGTGGACCA	138
Db			
OY	42	CysArGalaIleAlaProTyrPheThrGluLeuAlaLysAsnAsnProAsnValAlaPhe	61
		::	
Db	139	TGTGCTTTCACACGCTCCACTCTTGGTATTTGGCTTAAGAAACCTTCAACCTGCTTTTC	198
OY	62	LeuLysValAspValAspGluLeuAsnSerValAlaSerLysTrpGluIleAsnIleMet	81
Db	199	CTCAAGGTGTACTGATGATTAATTCAGTCGGGTGGCAATGATGGCGCATACAGCGATG	258
OY	82	ProThrPheValPheLeuLysLysGlyLysIleIleGluLysIleValAlaLysPlys	101
		::	
Db	259	CCAACCTTCACGTGTTTGGAGAAGAGGGAAGATTGTTGGACCAAGCCAAAGAA	318
OY	102	ValGlyLeuSerLysLysIle	108
Db	319	GATGAGCTTACGATCTACCATTT	339

RESULT	5
ATTHIOARA	
LOCUS	A <sup>+</sup> THIOARA                  497 bp       mRNA
DEFINITION	A. thaliana mRNA for thioredoxin H.
ACCESSION	Z11084
VERSION	Z11084.1    GI:16551
KEYWORDS	Thioredoxin H.
SOURCE	Arabidopsis thaliana (thale cress)
ORGANISM	Arabidopsis thaliana

REFERENCE  
AUTHORS  
TITLE  
JOURNAL  
MEDLINE  
PUBMED  
REFERENCE  
AUTHORS  
TITLE  
JOURNAL

1 (bases 1 to 497)  
Rivera-Medrid,R., Marinho,P., Brugidou,C., Chartier,Y. and Meyer,Y  
Nucleotide sequence of a cDNA clone encoding an Arabidopsis  
thaliana thiorodoxin h  
Plant Physiol. 102 (1), 327-328 (1993)  
94515431  
8108503  
2 (bases 1 to 497)  
Meyer,Y.  
Direct Submission  
Submitted (16-JUL-1992) Yves Meyer PhD, Laboratoire de Physiologie

FEATURES	Location/Qualifiers
source	1. .497

CDS 30.374

```

/product="Thioredoxin H"
/protein_id="CA78462.1"
/db_xref="GI:16552"
/db_xref="SWISS-PROT:P29448"
/translation="MASEEGVLIACHVETWNEDLQAKNEKTLVYVDFTASMGCPCT
FIAPFADLAKRLPNVFLFKVDITDELKSVASDMAIQAPFTFMLEKGIILDKVVGAKK
DELQSTIAKHIA"

```

BASE COUNT	135 a	94 c	116 g	152 t
ORIGIN				

Alignment Scores:	
Pred. No.:	2,04e-34
Score:	376.00
Percent Similarity:	77.57%
Best Local Similarity:	66.36%
Query Match:	59.03%
DB:	8
	Gaps: 0
	Indels: 0
	Mismatches: 24
	Conservative: 12
	Matches: 71
	Length: 497

US-09-786-715-2 (1-122) x ATTHIOARA (1-497)

Oy	2	llacuglugjyglvnaillalalcyshtslyslleaspglutrpglugllyglmeuonly	21
Db	36	TCGAGACAGGACACAGTATCGCTCCGCACACCGTTGAGACATGGAACGACGCTTCA	95
Oy	22	LysTrpLysAspSerGluLysLeuValValAlaSpheThrAlaSerTrpCysGlyPro	41
Db	96	AAGGCTATGATGCCAANAACCTCTGTGGTGGTGTTCACAGCGCTTTGGTGTGACCA	155
Oy	42	CysArgAlaIleAlaProTyrPheThrGluLeuAlaLysAsnAsnProAsnValAlaPhe	61
Db	156	TGTGGTTTCATCGCTCCATCTTTGCGATTTGGCTAAGAACTTCACAGCGTGGTTTC	215
Oy	62	LeuLysValaAspValaAspGluLeuAsnSerValAlaSerLysTrpGluIleAsnIleMet	81
Db	216	CTCAAGGTTGTACTGATGAATTAAGTCGGGTGGCAAGTGAATGGGCATGCAGGCGATG	275
Oy	82	ProThrPheValPheLeuLysLysGlyLysIleIleGluLysIleValGlyAlaAspLys	101
Db	276	CCAACCTTCATGTTTTTGAAGAAGGAGATTTTGGACAAGATTTGGAGGCCAAGAA	335
Oy	102	ValGlyLeuSerLysLysIle	108
Db	336	GATGAGCTTCAGTCTACCATTT	356

RESULT	6
RCTHIORXN	
LOCUS	603 bp.
DEFINITION	mRNA
ACCESSION	Z706877
VERSION	2706877.1 GI:1255953
KEYWORDS	thiredoxin.
SOURCE	Ricinus communis (castor bean)
ORGANISM	Ricinus communis
	linear PLN 04-APR-1996

REFERENCE	1 (bases 1 to 603)
AUTHORS	Szedekényi, J., Dolgenier, E. and Schobert, C.
TITLE	cdna expressed in Ricinus cotyledons
JOURNAL	unpublished
REFERENCE	2 (bases 1 to 603)
AUTHORS	Dolgenier, E.
TITLE	Direct Submission
JOURNAL	Submitted (04-ARR-1996) Dolgenier, E., University of Bayreuth, Department of Plant Physiology, Universitaetsstr.30, Bayreuth, Germany, 95440

source	Location/Qualifiers
1. .603	

```

/organism="Ricinus communis"
/mol_type="mRNA"
/cultivar="Sanguineus"
/db_xref="taxon:3988"
/clone="pEDRH018"
/tissue_type="cotyledon"
/clone_lib="lambda Excell"
/dev_stage="seedling"
23. .379

```

```

/product="thioredoxin"
/protein_id="CA94534.1"
/db_xref="GI:1255954"
/_db_xref="SPTREMBL:O43636"
/translation="MAAEAGVIGCHVEAANNEDLQKNDKGLIYDVFASMGCPGR
FIAFLEIAKLPENVFELKVDDELKTVAEHWEVSMPTFFLEKGRIMDKVGGARK
DELQGITAKIMATASn"

```

BASE COUNT	179	a	101	c	157	g	166	t
ORIGIN								

Alignment Scores:	
Pred. No.:	2,16e-33
Score:	368.00
Percent Similarity:	75.22%
Best Local Similarity:	60.18%
Length:	603
Matches:	68
Conservative:	17
Mismatches:	28

Query Match: 57.77% Indels: 0  
 DB: 8 Gaps: 0  
 US-09-786-715-2 (1-122) x RCHIORXN (1-603)

OY 2 AlaGluGluGluValIleAlaCysHisIleAspGluTrpGluGluLeu 21  
 DB 29 GCAGAGAAAGGCAAGTATCGGTCACACTGTGAGCATGAGCAATTCGAG 88  
 OY 22 LysTrpLysAspSerGluLysLeuValValAlaAspPheThrAlaSerTrpCysGly 41  
 DB 89 AAGGAAATGATACCAAGGAGTATCGTTGATTATTCGCTTCATGAGTGCACCA 148  
 OY 42 CysArgAlaIleAlaProTyrPheThrGluLeuAlaLysAsnProAsnValAlaPhe 61  
 DB 149 TGCCCTTCATTCGCTCCCTTCCTTGGCTGAGCTGCCAAGAACTCCCAATGTTACCTTC 208  
 OY 62 LeuLysValAspValAspGluLeuAsnSerValAlaSerLysTrpGluIleAsnAlaMet 81  
 DB 209 CTGAAGTGTGATGTGATGAACTGAAAGACTGTGCTCAAGAGTGGCTGTGAGTCATAG 268  
 OY 82 ProThrPheValPheLeuLysGlyLysIleIleGluLysIleValGlyAlaAspLys 101  
 DB 269 CCACCTTATGCTTCCTGAGAGGGAAGATATGACAAAGCTGTTGAGCAAGAA 328  
 OY 102 ValGlyLeuSerLysIleLeuGluLeuSerGlyThr 114  
 DB 329 GACGACTGCACAACTATACGCAACACATGCTACT 367

RESULT 7  
 LOCUS AY170650 665 bp mRNA linear PLN 02-JAN-2003  
 DEFINITION Pisum sativum clone 3 thioedoxin h mRNA, complete cds.  
 ACCESSION AY170650  
 VERSION AY170650.1 GI:27466893  
 KEYWORDS  
 SOURCE  
 ORGANISM

Pisum sativum (pea)  
 Pisum sativum  
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
 rosids; eurosids I; Fabales; Fabaceae; Papilionoideae; Viciae;  
 Pisum.

REFERENCE  
 AUTHORS Montichard, F., Renard, M., Duval, F.D. and Macherel, D.  
 TITLE Expression of the NADP-thioedoxin reductase/thioedoxins h system during germination of seeds of Pisum sativum L  
 JOURNAL Unpublished  
 REFERENCE 2 (bases 1 to 665)  
 AUTHORS Montichard, F., Renard, M., Duval, F.D. and Macherel, D.  
 TITLE Direct Submission  
 JOURNAL Submitted (30-OCT-2002) UMR 1191 Physiologie Moléculaire des Semences, UFR Sciences, 16 bd Lavoisier, Angers 49045, France

FEATURES  
 location/Qualifiers  
 1..665  
 /organism="Pisum sativum"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:3888"  
 /clone="3"  
 74..415  
 /note="ex h"  
 /codon\_start=1  
 /product="Chloredoxin h"  
 /protein\_id="AA012854.1"  
 /db\_xref="GI:27466894"  
 /translation="MAEGQYIVHTDAMKEQLEKASKKLIYDFTASMGPCRF  
 IAPILAIATKLTHTFLKVDVDELKTVSEMGIEAMPTFLKNGELVDKVGAKKE  
 ELQKIDKHAA"

## CDS

BASE COUNT 220 a 85 c 164 g 196 t  
 ORIGIN

Alignment Scores:  
 Pred. No.: 354.00  
 Score: 67

Percent Similarity: 75.93% Conservative: 15  
 Best Local Similarity: 62.04% Mismatches: 26  
 Query Match: 55.57% Indels: 0  
 DB: 8 Gaps: 0  
 US-09-786-715-2 (1-122) x AY170650 (1-665)

OY 1 MetAlaGluGluGluValIleAlaCysHisIleAspGluTrpGluGluLeu 20  
 DB 74 ATGGCGGAGAGGAGCAAGTATCGGTCACACCGTGGATGCTGGAGAGACAGTA 133  
 OY 21 GlyLysTrpLysAspSerGluLysLeuValValAlaAspPheThrAlaSerTrpCysGly 40  
 DB 134 GAGAAAGGGAAGCCCTCAAGAAAGTATGATGATTTACGCTCTTGCGCGGT 193  
 OY 41 ProCysArgAlaIleAlaProTyrPheThrGluLeuAlaLysAsnProAsnValAla 60  
 DB 194 CCATCCCGTTTATTCGCCCAATTTTGGCAGAGATTTGTAACACATGCTACT 253  
 OY 61 PheLeuLysValAspValAspGluLeuAsnSerValAlaSerLysTrpGluIleAsnAla 80  
 DB 254 TTCCTTAAGTGTGAGTGTGATGAAAGACTGTTCCGAGAGTGGGGAATTCAGCT 313  
 OY 81 MetProThrPheValPheLeuLysGlyLysIleIleGluLysIleValGlyAlaAsp 100  
 DB 314 ATGCCAATCTTCTGCTGGAAGATGTGTAACCTTGAGCAAGTGTGGTCCCAAG 373  
 OY 101 LysValGlyLeuSerLysIle 108  
 DB 374 AAGGAGAGCTGCATTCATTCGAAAT 397

RESULT 8  
 LOCUS D87984 595 bp mRNA linear PLN 07-FEB-1999  
 DEFINITION Fagopyrum esculentum mRNA for thioedoxin, complete cds.  
 ACCESSION D87984  
 VERSION D87984.1 GI:1620904  
 KEYWORDS  
 SOURCE  
 ORGANISM

Fagopyrum esculentum (common buckwheat)  
 Fagopyrum esculentum  
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
 Caryophyllales; Polygonaceae; Fagopyrum.

REFERENCE  
 AUTHORS Fujino, K., Funatsuki, H. and Kikuta, Y.  
 TITLE buckwheat cDNA from immature seeds  
 JOURNAL Unpublished  
 REFERENCE 2 (bases 1 to 595)  
 AUTHORS Fujino, K.  
 TITLE Direct Submission  
 JOURNAL Submitted (24-SEP-1996) Kaizen Fujino, Hokkaido University, Faculty of Agriculture, Kita 9 Nishi 9, Kita-ku, Sapporo, Hokkaido 060, Japan (E-mail: kaizen@2.hines.hokudai.ac.jp, Tel: 011-706-2482, Fax: 011-706-2471)

FEATURES  
 location/Qualifiers  
 1..595  
 /organism="Fagopyrum esculentum"  
 /mol\_type="mRNA"  
 /strain="cv. Kitayuki"  
 /db\_xref="taxon:3617"  
 86..436  
 /codon\_start=1  
 /product="thioedoxin"  
 /protein\_id="BA13524.1"  
 /db\_xref="GI:1620905"  
 /translation="MAEAGYIACHTVQENNEKFOKADSGKLIYDFTASMGPCRF  
 ITPYSELAKKEPPIVAFKVVDDLDKVAEEYKVEAMSPFVLKGGQVERIVGAKRD  
 ELIKIVAHVITA"

## CDS

BASE COUNT 170 a 107 c 157 g 161 t  
 ORIGIN

Alignment Scores:  
 Pred. No.: 1.51e-31  
 Length: 595

Score: 352.00 Matches: 66  
 Percent Similarity: 75.93% Conservative: 16  
 Best Local Similarity: 61.11% Mismatches: 26  
 Query Match: 55.26% Indels: 0  
 DB: 8 Gaps: 0

US-09-786-715-2 (1-122) x D87984 (1-595)

QY 1 MetAlaGluGluGlyGlnValIleAlaCysHisLysLysLeuAspGluTrpGluGlnLeu 20  
 DB 86 ATGGCGAAGACACACAGATTATGCGATGCGACACCGGTCCAGAGTGGAAAGCAAGTTC 145  
 QY 21 GLYSTPLYSASPSERGIULYSLEUValValAlaAspPheThrAlaSerTPCYsGly 40  
 DB 146 CAGAAAGCCAGATTCCCGCAATTCATGTCATGATTCATCCAGCATATGTCGGT 205  
 QY 41 ProCysArgAlaIleAlaProTyrPheThrGluLeuAlaLysAsnProAsnValAla 60  
 DB 206 CCTCGCGCGCTCAATCACTCCGTACGTCTGACCTGCCAAGAGTTTCCTCATGTGGCC 265  
 QY 61 PheLeuLysValAlaAspGluLeuAsnSerValAlaSerLysTrpGluIleAsnAla 80  
 DB 266 TTCTTCAAGTGAATGTTGATTTGAAGATGTCCTCAGGATACAGAGTGGAGCA 325  
 QY 81 MetProThrPheValPheLeuLysGlyLysIleIleGluLysIleValAlaAsp 100  
 DB 326 ATGCCGAGCTTGTATTACTGAAGAAGGCGCAGAGAGTTGAAAGATTGTGGGTCCAGG 385  
 QY 101 LysValGlyLeuSerLysLysIle 108  
 DB 386 AAGGACGAACCTCCTCATAGATA 409

# RESULT 9

PSA310990 630 bp mRNA linear PLN 12-APR-2001  
 LOCUS Pisum sativum mRNA for thioredoxin h (trxn gene).  
 DEFINITION AJ310990.1 GI:13624883  
 VERSION thioredoxin H; trxn gene.  
 KEYWORDS Pisum sativum (pea)  
 SOURCE Pisum sativum  
 ORGANISM Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids; eurosids I; Fabales; Fabaceae; Papilionoideae; Viciae;

REFERENCE 1  
 AUTHORS Traverso, J.A.  
 JOURNAL Thesis (2001). Department of Bioquímica, Biología Celular y Molecular de Plantas, Estación Experimental del Zaidin (C.S.I.C.), Granada, Spain

REFERENCE 2  
 AUTHORS Traverso, J.A., Cazalis, R., Sahrawy, M., Lopez-Gorge, J. and Chueca, A.  
 JOURNAL Sequencing, cloning and expression of Pisum sativum thioredoxin h (bases 1 to 630)  
 REFERENCE Traverso, J.A.  
 TITLE Direct Submission  
 JOURNAL Submitted (11-APR-2001) Traverso J.A., Biología Celular y Molecular de plantas, Estación Experimental del Zaidin (CSIC), Profesor Albareda 1, Granada, 18008, SPAIN

FEATURES  
 source 1. 630  
 /organism="Pisum sativum"  
 /mol\_type="mRNA"  
 /variety="Lincoln"  
 /db\_xref="taxon:3888"  
 /tissue\_type="Leaves"  
 1. 630  
 /gene="trxn"  
 36. 398  
 /gene="trxn"  
 /function="Unknown"  
 /codon\_start=1

/evidence=experimental  
 /product="thioredoxin h"  
 /protein\_id="CAC36986.1"  
 /db\_xref="GI:13624884"  
 /translation="MAGSSEGVISCHSDANDNLIHRNEKSLIVDFTRSKCP  
 CRFAPFLGLAKFTNVIPLKVDLSKADMDVMAEAPTFVKEESTILGKVGGA  
 KDELQOTIERHVAASSNA"  
 398. .609  
 /gene="trxn"  
 /gene="trxn"

3' UTR  
 polyA\_site

BASE COUNT 195 a 95 c 141 g 199 t

ORIGIN

Alignment Scores:  
 Pred. No.: 1 61e-31 Length: 630  
 Score: 352.00 Matches: 66  
 Percent Similarity: 74.77% Conservative: 14  
 Best Local Similarity: 61.68% Mismatches: 27  
 Query Match: 55.26% Indels: 0  
 DB: 8 Gaps: 0

US-09-786-715-2 (1-122) x PSA310990 (1-630)

QY 2 AlaGluGluGlyGlnValIleAlaCysHisLysLysLeuAspGluTrpGluGlnLeuGly 21  
 DB 48 TCAGAGAGGAGACAGATTATCAGCTGCACAGCGTTCATGAGAACGATATTCCTCAC 107  
 QY 22 LysTrpLysAspSerGIULYSLEUValValAlaAspPheThrAlaSerTPCYsGlyPro 41  
 DB 108 AGAGGCATATCAATCAAGAAACGATGTTGTGACCTTCCTCTTGTGTGGACCA 167  
 QY 42 CysArgAlaIleAlaProTyrPheThrGluLeuAlaLysAsnProAsnValAlaPhe 61  
 DB 168 TGCCGTTTCATTGACCATTCCTTGGAATTCGCTAAGAAGTTACAAATGTCATTTTC 227  
 QY 62 LeuLysValAlaAspValAlaAspGluLeuAsnSerValAlaSerLysTrpGluIleAsnAlaMet 81  
 DB 228 CTTAAGGTGACGCGAGCAAGTAAGCTGTTGCTCAGATGTTGGCTGTGAGGCTATG 287  
 QY 82 ProThrPheValPheLeuLysGlyLysIleIleGluLysIleValAlaAspLys 101  
 DB 288 CCAACATTTGTTGTGAAAGAAAGCAATTTGGCAAGTGTGGAGCAAAAGAAA 347  
 QY 102 ValGlyLeuSerLysLysIle 108  
 DB 348 GAGGAACGCGACGACGACATTT 368

# RESULT 10

LOCUS NTTRNA 698 bp mRNA linear PLN 15-FEB-1994  
 DEFINITION N. tabacum mRNA for thioredoxin.  
 ACCESSION X58527.1 GI:20046  
 VERSION X58527.1 GI:20046  
 KEYWORDS Nicotiana tabacum (common tobacco)  
 SOURCE Nicotiana tabacum  
 ORGANISM Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Asteridae; lamids; Solanales; Solanaceae; Nicotiana.

REFERENCE 1  
 AUTHORS Marty, I. and Meyer, Y.  
 JOURNAL Nucleotide sequence of a cDNA encoding a tobacco thioredoxin (bases 1 to 698)  
 MEDLINE Plant Mol. Biol. 17 (1), 143-147 (1991)  
 PUBMED 91328721

REFERENCE 2  
 AUTHORS Brugidou, C., Marty, I., Chartier, Y. and Meyer, Y.  
 JOURNAL The Nicotiana tabacum genome encodes two cytoplasmic thioredoxin genes which are differentially expressed  
 MEDLINE Mol. Gen. Genet. 238 (1-2), 285-293 (1993)  
 PUBMED 93241165  
 8479434

REFERENCE 3 (bases 1 to 698)  
 AUTHORS Meyer, Y.  
 TITLE Direct Submission  
 JOURNAL Submitted (14-FEB-1991) Y. Meyer, Laboratoire de Physiologie et Biologie, Moleculaire Vegetale, Universite av de Villeneuve, 66860 Perpignan, France  
 COMMENT Gene product is probably cytoplasmic.  
 FEATURES  
 source  
 location/Qualifiers  
 1..698  
 /organism="Nicotiana tabacum"  
 /mol\_type="mRNA"  
 /strain="white Burley"  
 /db\_xref="taxon:4097"  
 /clone\_id="plasmid PT219"  
 /dev\_stage="in vitro cells"  
 1..682  
 /evidence=experimental  
 79..459  
 /codon\_start=1  
 /product="thioredoxin"  
 /protein\_id="CA41415.1"  
 /db\_xref="GI:20047"  
 /db\_xref="SWISS-PROT:P29449"  
 /translation="MAANDATSEGGVFGCHKVENNEYFKGVETKLVVDFTAS  
 WCGPRTIADIAIKKMPHVFILKVDDELKTVSAEWSVEAMPTEVFITKDKVEDR  
 VVGAKKEELQGTIVKHAAPATVTA"  
 polyA\_signal  
 468..473  
 polyA\_signal  
 511..517  
 polyA\_site  
 682  
 BASE COUNT 192 a 117 c 164 g 225 t  
 ORIGIN  
 Alignment Scores:  
 Pred. No.: 4.62e-31 Length: 698  
 Score: 348.50 Matches: 65  
 Percent Similarity: 72.73% Conservative: 23  
 Best Local Similarity: 53.72% Mismatches: 30  
 Query Match: 54.71% Indels: 3  
 Gaps: 1  
 US-09-786-715-2 (1-122) x NTTRNA (1-698)  
 QY 2 ALaGlUGlUGlGlnValIleAlaCySHsLysIleAspGluTrpGluGlnLeuGly 21  
 Db 103 TCCGAGAGGAGCAAGTGGCTGGCCACAAAGTTGAGCATGGAACGAGTACTTCAAG 162  
 QY 22 LysTrpLysAspSerGluLysLeuValValAlaValAspPheThrAlaSerTTPCysGlyPro 41  
 Db 163 AAAGCGCTTGAGACTAGAAACTGGTGGTGCATTTTACTGCTCATGGTGGCGCCCT 222  
 QY 42 CysArgAlaIleAlaProTyrPheThrGluLeuAlaLysAsnAspProAsnValAlaPhe 61  
 Db 223 TGGCGTTTATTGGCCCAATTTCTGTCGACTGCTAAGAAAGATGCCCATGTAATATTC 282  
 QY 62 LeuLysValAspValAspGluLeuAsnSerValAlaSerLysTrpGluIleAsnAlaMet 81  
 Db 283 CTCAGAGTTGATGTGTGATGACACTGACACTGTTTACCGGATGAGATGGAGCGCAATG 342  
 QY 82 ProThrPheValPheLeuLysLysGlyLysIleIleGluLysIleValGlyAlaAspLys 101  
 Db 343 CCAACTTTTCTTCATTAAGATGAAGAAAGATGACAGACTGTTGGTGGCCAGAAA 402  
 QY 102 ValGlyLeuSerLysLysIleLeuGluLeuSerGlyThrThrProAlaIleThrSerThr 121  
 Db 403 GAGGAGTTGCAGCAGACCATGATGACATGCTGCTCTCTCTCTCTCTCTCTCTCTCTCT 453  
 QY 122 Ala 122  
 Db 454 GCT 456  
 RESULT 11  
 LOCUS AY088698 590 bp mRNA linear PLN 14-APR-2003

DEFINITION Arabidopsis thaliana clone 9219 mRNA, complete sequence.  
 ACCESSION AY088698  
 VERSION AY088698.1 GI:21407472  
 KEYWORDS FLI\_CDNA.  
 SOURCE Arabidopsis thaliana (thale cress)  
 ORGANISM Arabidopsis thaliana  
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsids.  
 1 (bases 1 to 590)  
 Haas, B.J., Volfovsky, N., Town, C.D., Troupkhan, M., Alexandrov, N., Feldmann, K.A., Flavell, R.B., White, O. and Salzberg, S.L.  
 Full-length messenger RNA sequences greatly improve genome annotation  
 Genome Biol. 3 (6), RESEARCH0029 (2002)  
 JOURNAL MEDLINE 22088475  
 PUBMED 12093376  
 REFERENCE 2 (bases 1 to 590)  
 AUTHORS Brover, V., Troupkhan, M., Alexandrov, N., Lu, Y.-P., Flavell, R. and Feldmann, K.  
 TITLE Full-length cDNA from Arabidopsis thaliana unpublished  
 JOURNAL 3 (bases 1 to 590)  
 REFERENCE Brover, V., Troupkhan, M., Alexandrov, N., Lu, Y.-P., Flavell, R. and Feldmann, K.  
 TITLE Direct Submission  
 JOURNAL Submitted (11-MAR-2002) Ceres, Inc, 3007 Malibu Canyon Road, Malibu, CA 90265, USA  
 COMMENT This clone sequence is one of 5,000 Ceres full-length cDNAs made available to TIGR and Genbank. The following quality assessment of this set was done by comparison with known proteins: less than one percent are 3'-truncated; approximately two percent represent alternative splice variants, including unspliced introns and spliced exons; one percent may contain premature stop codons; five percent may have frame shifts in a coding region. A sequence is considered to be 5'-truncated if it lacks the translation initiation start (ATG). A sequence is considered to be 3'-truncated if it lacks the C-terminal end of the encoded protein. Please note that these cDNA sequences are derived from the Ws or Wier ecotypes and therefore may contain polymorphisms when compared to sequences from Col-0. Genset carried out the library production and sequencing of the full-length clones. Ceres, Inc. carried out the clustering of the 5' sequences, selection of clones, and sequence assembly.  
 FEATURES  
 source  
 location/Qualifiers  
 1..590  
 /organism="Arabidopsis thaliana"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:3702"  
 /clone="9219"  
 52..411  
 /codon\_start=1  
 /product="thioredoxin"  
 /protein\_id="AA067018.1"  
 /db\_xref="GI:21617968"  
 /translation="MAAEQGVNSCHNDVYQVOLDKAKESNKLIIVDFASWCPQR  
 MIRIFNDLAKKEMSAIFRKVDVDELQSVAKKEGVAMPTFVFIKAGEYVDKLVGAN  
 KEDLQAKIVHGTGVT"  
 BASE COUNT 173 a 92 c 142 g 183 t  
 ORIGIN  
 Alignment Scores:  
 Pred. No.: 6.48e-31 Length: 590  
 Score: 346.50 Matches: 68  
 Percent Similarity: 74.36% Conservative: 19  
 Best Local Similarity: 58.12% Mismatches: 29  
 Query Match: 54.40% Indels: 1  
 Gaps: 1  
 US-09-786-715-2 (1-122) x AY088698 (1-590)  
 QY 2 ALaGlUGlUGlGlnValIleAlaCySHsLysIleAspGluTrpGluGlnLeuGly 21  
 ||||||||||||||||||||||||||||| ||| ||| |||||||





[illegible]

**COMMENT**

230-0045, Japan (E-mail:msekigsc.riken.go.jp,  
URL:<http://pfweb.gsc.riken.go.jp/>,  
Fax:81-45-503-9586)  
An Arabidopsis full-length cDNA library was constructed essentially  
as reported previously [Seki et al., (1998) Plant J. 15:707-720;  
Seki et al., (2002) Science 296:141-145]. cDNA cleaved with BamHI  
and XhoI was ligated to modified Lambda FLX-I-E vector (carinci et  
al., (2001) Genomics 77:79-90) digested with BamHI and SalI.  
This clone is in a modified pluscript vector.  
Please visit our web site (<http://pfweb.gsc.riken.go.jp/>) for  
further details.

**FEATURES**

source	location/Qualifiers
	1..556
	/organism="Arabidopsis thaliana"
	/mol_type="mRNA"
	/cultivar="Columbia"
	/db_xref="taxon:3702"
	/chromosome="1"
	/clone="RAF19-77-A10"
	/note="common name: thale cress"
gene	1..556
	/gene="Atlg19730/F1APL_32"
CDS	23..382
	/gene="Atlg19730/F1APL_32"
	/codon_start=1
	/product="putative thioredoxin"
	/protein_id="BAC43145.1"
	/db_xref="gi:26452111"
	/translation="MAAEQGVICGHFNVDYVTLQDKAKESNKLIVIDFTSMCPDPCKE MIAPFNDIAKKEMSSAIPFFVDVDELQSAKKEGVEAMPFPFIKAGEVVDDKLGNAN KEDIQARKTGYTTTA"
BASE COUNT	169 a      88 c    132 g     167 t
ORIGIN	
Predicted Scores:	
Align. No.:	7.9e-31                  length:        556
Score:	345.50                  Matches:        68
Best Local Similarity:	73.50%                  Conservative:    18
Query Match:	58.12%                  Mismatches:     30
	54.24%                  Indels:           1
DB:	Gaps:                    1

```
US-09-786-715-2 (1-122) x AK118542 (1-556)
```

OY	2	AAlaglugluguylglnvallllealaCyshstlsySlleaSpGlutrgplgllyGlnleugly	21
Db	29	GCACAAGGGGTCAAGTGATGGTTGTTCACAGAACAAGCATGATGACTGTCCAACTTAT	88
OY	22	LysTrrpLySaSPserGLyLSLeuValVaIalaspherhralaserTrpCYsgLYPro	41
Db	89	AAAACC AAAAGAACATCCAACACGCGATTGTGTGA TTTCAC TCGTTCATCGTGTGCCACCA	148
OY	42	CysArghAllllealaprotyrPhetheThglueuleualysAsnaSnproASnValALA---	60
Db	149	TGCCGCGATGATGGTCTCCATTTC AAC CATT TGCC CA A G ACT TC AT CGTCAA GT GCC AT C	208
OY	61	pHeuenuYSlaSPvaIaspGIluenuasnSerValaIseryLTsrPGullleasnAla	80
Db	209	TTCTTCAAGGTGATGTGATGACACTTCAGAGTGTGCTTAANAAGTTTGGTGTGGAGCA	268
OY	81	MetProThrPheValpHeleunylslGLySLlysIlleillegulySIlevalGIyaLaASP	100
Db	269	ATGCGAACCTTTGTGTTTCATTAAAGCCGCGAAMTGGTAGTGAAGCTGCTGTGGCGAAT	328
OY	101	LysValGIyseurLyslySIllaecGUeneuseerGLythrPrroALa	117
Db	329	AAAA GA AG ATT CTTCAG GC GA AA A V TA GTA GA CAT AC TG GT TTA CA ACA CTC GG	379

RESULT 15

LOCUS	AF051206	641 bp	mRNA	linear	PLN 18-AUG-1998
DEFINITION	Picea edulis probable thioredoxin H (SD09) mRNA, complete cds.				
ACCESSION	AF051206				



GenCore version 5.1.6  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - nucleic search, using frame\_plus\_p2n model

Run on: August 17, 2003, 17:46:53 ; Search time 171.507 Seconds

(without alignments)  
1920.219 Million cell updates/sec

Title: US-09-786-715-2

Perfect score: 637

Sequence: 1 MAEEGOVYACHKIDEMEQQL.....GLSKITELISGTTPATSTVA 122

Scoring table:

BLOSUM62			
Xgapop 10.0	Xgapext 0.5		
Ygapop 10.0	Ygapext 0.5		
Fgapop 6.0	Fgapext 7.0		
Delop 6.0	Delext 7.0		

Searched: 2552756 seqs, 1349719017 residues

Total number of hits satisfying chosen parameters: 5105512

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame\_p2n.model -DEV=xlp  
-o=/cgn2.1/USPTO/US09786715/runat\_11082003\_150513.6025/app\_query.fasta\_1.1052  
-DB=N\_Geneseq\_19Jun03 -QFMT=fastap -SUFFIX=ring -MINMATCH=0.1 -LOOPCL=0  
-LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi  
-LIST=45 -DOCALLIGN=200 -THR.SCORE=pct -THR.MAX=100 -THR.MIN=0 -ALIGN=15  
-MODE=LOCAL -OUTFMT=ptio -NORR=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000  
-USER=US09786715.ecgn.1.1.874@runat\_11082003\_150513.6025 -NCPD=6 -ICPU=3  
-NO\_MMAP -LARGEQUERY -NEG.SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG  
-NO.TIMEOUT=120 -WARN.TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6  
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

N\_Geneseq\_19Jun03:\*

1:	/SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1980.DAT:*
2:	/SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1981.DAT:*
3:	/SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1982.DAT:*
4:	/SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1983.DAT:*
5:	/SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1984.DAT:*
6:	/SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1985.DAT:*
7:	/SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1986.DAT:*
8:	/SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1987.DAT:*
9:	/SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1988.DAT:*
10:	/SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1989.DAT:*
11:	/SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1990.DAT:*
12:	/SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1991.DAT:*
13:	/SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1992.DAT:*
14:	/SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1993.DAT:*
15:	/SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1994.DAT:*
16:	/SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1995.DAT:*
17:	/SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1996.DAT:*
18:	/SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1997.DAT:*
19:	/SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1998.DAT:*
20:	/SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1999.DAT:*
21:	/SIDSI/gcgdata/geneseq/geneseqn-emb1/NA2000.DAT:*
22:	/SIDSI/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT:*
23:	/SIDSI/gcgdata/geneseq/geneseqn-emb1/NA2001B.DAT:*
24:	/SIDSI/gcgdata/geneseq/geneseqn-emb1/NA2002.DAT:*
25:	/SIDSI/gcgdata/geneseq/geneseqn-emb1/NA2003.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed,

and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	637	100.0	870	21	AAZ51737	Momordica charantia
2	377.5	59.3	509	22	AAH87768	Peppermint plant o
3	377	59.2	3888	24	ABSS3097	DNA encoding Thlor
4	377	59.2	3888	24	ABN89581	Phaseolin promoter
5	376.5	59.1	470	21	AAC37781	Arabidopsis thalia
6	376	59.0	445	24	AB212358	Arabidopsis thalia
7	376	59.0	345	24	ABN89580	Arabidopsis thalia
8	376	59.0	561	21	AAC51522	Arabidopsis thalia
9	376	59.0	563	21	AAC34121	Arabidopsis thalia
10	376	59.0	3129	24	ABSS3095	DNA encoding thlor
11	376	59.0	3129	24	ABN89579	Phaseolin promoter
12	376	59.0	3888	24	ABSS3096	DNA encoding Oleos
13	376	59.0	3888	24	ABN89580	Phaseolin promoter
14	375	58.9	4935	24	ABN89586	Promoter-Oleoin t
15	366	57.5	738	21	AAZ51739	Glycine max thior
16	361	56.7	574	21	AAZ51738	Catalpa speciosa t
17	360.5	56.6	402	25	ABX21664	Human GDP-mannose
18	357	56.0	328	21	AAA31097	Plant microsateili
19	354	55.6	392	25	ABX18045	Human GDP-mannose
20	354	55.6	601	21	AAZ51740	Glycine max thior
21	352	55.3	346	21	AAA31096	Plant microsateili
22	350	54.9	614	21	AAZ51741	Vernonia mespilifo
23	348	54.6	320	21	AAA31785	Plant microsateili
24	346.5	54.4	590	21	AAC38792	Arabidopsis thalia
25	345.5	54.2	360	24	ABZ12359	Arabidopsis thalia
26	343	53.8	390	25	ABX19403	Human GDP-mannose
27	343	53.8	686	15	AAO78205	Gene coding for pr
28	343	53.8	687	21	AAC66375	Rice thiodoxin h
29	337	52.9	560	21	AAC41961	Arabidopsis thalia
30	335.5	52.7	653	16	AAO99783	Plant SAR gene pl.
31	335.5	52.7	653	20	AAV62799	Tobacco SAR CHX in
32	335.5	52.7	653	20	AAV81683	Tobacco protein-sy
33	335	51.0	419	21	AAA31942	Plant microsateili
34	324.5	50.9	524	21	AAC33829	Arabidopsis thalia
35	324	50.9	292	21	AAA31978	Plant microsateili
36	323	50.7	357	24	ABZ13931	Arabidopsis thalia
37	323	50.7	480	21	AAC36542	Arabidopsis thalia
38	323	50.7	563	21	AAC34211	Arabidopsis thalia
39	323	50.7	562	21	AAC48656	Arabidopsis thalia
40	315	49.5	572	21	AAC52069	Arabidopsis thalia
41	315	49.5	576	25	ABX56868	Arabidopsis thalia
42	314	49.3	353	21	AAA31118	Plant microsateili
43	312	49.0	328	21	AAA31222	Plant microsateili
44	311	48.8	274	21	AAA31859	Plant microsateili
45	310	48.7	356	21	AAA31804	Plant microsateili

#### ALIGNMENTS

RESULT 1  
AAZ51737  
ID AAZ51737 standard; cDNA: 870 BP.  
AAZ51737:  
04-JUL-2000 (first entry)  
Momordica charantia thiodoxin cDNA.  
Momordica charantia thiodoxin: clone fds.pk0001.e9: chimeric gene;  
transgenic plant; seed storage protein; allergenicity; ss.  
Momordica charantia.  
Key Location/Qualifiers  
CDS 49..417

```

FT      /*tag= a
XX      /product= "thioredoxin"
XX      WO200014239-A2.
XX      16-MAR-2000.
XX      07-SEP-1999; 99WO-US20420.
XX      08-SEP-1998; 98US-0099501.
XX      (DUPO ) DU FONT DE NEMOURS & CO E I.
XX      Allen SM, Thorpe CJ, Lu AL;
XX      WPI; 2000-256987/22.
XX      P-PSDB; AAY70479.
XX      New isolated polynucleotide encoding thioredoxin polypeptide is useful
XX      for producing transgenic plants with an altered level of thioredoxin -
XX      Claim 3; Page 28; 33pp; English.
XX      The present cDNA sequence encodes Momordica charantia thioredoxin
XX      protein. The cDNA was derived from clone fds-PK0001.e9, which was
XX      isolated from a cDNA library prepared from M. charantia developing seed
XX      tissue. Chimeric genes encoding all or a portion of the thioredoxin
XX      protein, in sense or antisense orientation are constructed, wherein
XX      expression of the chimeric gene results in production of altered levels
XX      of the thioredoxin protein in a transformed host cell. Thioredoxin is
XX      involved in the disassembly of seed storage proteins during germination
XX      by reducing S-S bonds and in the bread making process. Over expression of
XX      thioredoxin in cereals may reduce the allergenicity of any transgenic
XX      protein engineered into cereal crops with high sulfhydryl content.
XX      Sequence 870 BP; 236 A; 135 C; 209 G; 290 T; 0 other:

```

```

Alignment Scores:
Pred. No.: 5.9e-76 Length: 870
Score: 637.00 Matches: 122
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 21 Gaps: 0
US-09-786-715-2 (1-122) x AA251737 (1-870)

```

```

OY      1 MetAlaGluGluGluGluValIleAlaCysHisLysIleAspGluTrpGluGluGluLeu 20
DB      49 ATGCGTGAACAGAGGCGAAGTGGTGGTGCATTAAGATGAATGGAGGCAATTA 108
OY      21 GlyLysTrpLysAspSerGluLysLeuValValAspPheThrAlaSerTrpCysGly 40
DB      109 GGAATAATGGAAGATTCGTGAACACTGGTGTGGATTTTACTGCTCTCTGTCGGG 168
OY      41 ProCysArgAlaIleAlaProTyrPheThrGluLeuAlaLysAsnAsnProAsnValAla 60
DB      169 CCATGCCGGGCAATTCCTCATTTTCACGAATTCGCTAAGATTAACCAATGTCGCT 228
OY      61 PheLeuLysValAspValAspGluLeuAsnSerValAlaSerLysTrpGluIleAsnAla 80
DB      229 TTCTCTAAAGTCACGCTGACGAAATGAACAGTGTGCTAAGCAAGTGGGATTAATGCA 288
OY      81 MetProThrPheValPheLeuLysLysGlyLysIleIleGluLysIleValGlyAlaAsp 100
DB      289 ATGCCAAGCTTTTTCCTCGAATAAAAGGAAATAATTAAGAGATCGTTGCTGCTGAT 348
OY      101 LysValGlyLeuSerLysLysIleLeuGluLeuSerGlyThrThrProAlaIleThrSer 120
DB      349 AAAGTGGGCTGTGCAAGAAATATTAAGAGCTTAAGGAACCTACCTCCGCTGCTACTTC 408
OY      121 ThrAla 122
DB      111111

```

```

DB      409 ACTGCT 414
RESULT 2
AAH87768
ID      AAH87768 standard; cDNA; 509 BP.
XX      AAH87768;
AC      25-SEP-2001 (first entry)
DT      25-SEP-2001 (first entry)
XX      Peppermint plant oil gland expressed cDNA 124.
DE      Peppermint; plant oil gland cell; terpenoid essential oil; resin;
XX      genetic mapping; antisense suppression; recombinant expression; ss.
XX      Mentha x piperita.
XX      WO200153319-A1.
PN      26-JUL-2001.
PD      19-JAN-2001; 2001WO-US02567.
XX      20-JAN-2000; 2000US-0177264.
PR      (CROT/) CROTEAU R B.
XX      (LANG/) LANGE B M.
PA      (WILD/) WILDUNG M R.
XX      Croteau RB, Lange BM, Wildung MR;
PI      WPI; 2001-488706/53.
XX      New nucleic acid molecules corresponding to mRNA molecules expressed in
XX      peppermint oil glands for enhancing expression of plant oil gland cell
XX      proteins -
XX      Claim 1; Page 125; 251pp; English.
XX      The invention relates to nucleic acid molecules (AAH87645-AAH88116) that
XX      correspond to all or part of a mRNA molecule expressed in plant oil
XX      gland cells, especially peppermint and plant oil glands that produce
XX      terpenoid essential oils and resins. The nucleic acids are useful for
XX      genetically mapping a plant genome for genes expressed in plant oil
XX      gland cells and to suppress (for example by antisense suppression) or
XX      enhance their expression (for example by genetically transforming a
XX      plant cell with a replicable expression vector that expresses one or more
XX      proteins naturally expressed in plant oil gland cells). The nucleic acids
XX      are also useful for recombinant expression of plant oil gland proteins
XX      required for terpenoid essential oil and/or resin production in bacterial
XX      and/or yeast cells.
XX      Sequence 509 BP; 144 A; 103 C; 129 G; 133 T; 0 other:
US-09-786-715-2 (1-122) x AAH87768 (1-509)
Alignment Scores:
Pred. No.: 2.71e-41 Length: 509
Score: 377.50 Matches: 77
Percent Similarity: 73.33% Conservative: 11
Best Local Similarity: 64.17% Mismatches: 28
Query Match: 59.26% Indels: 4
DB: 22 Gaps: 1
OY      4 GluGlyGluValIleAlaCysHisLysIleAspGluTrpGluGluGluLysTrp 23
DB      50 GAAGGACAGGTGATCGGCTGCCACCACTGAATCACTGAACAGAGAGCTTCAGAGGCG 109
OY      24 LysAspSerGluLysLeuValValAspPheThrAlaSerTrpCysGlyProCysArg 43
DB      110 AATGATACACAGAGATTGGTAGTGGATTTTCACTGCTTCCGCTGCGGCTTGTCCG 169
OY      44 AlaIleAlaProTyrPheThrGluLeuAlaLysAsnAsnProAsnValAlaPheLeuLys 63

```



XX 19-DEC-2001; 2001WO-US50240.  
XX 19-DEC-2000; 2000US-0742900.  
PR 05-JUL-2001; 2001US-302885P.  
PR 04-DEC-2001; 2001US-0006038.  
XX (SEMB-) SEMBIOSIS GENETICS INC.  
XX (SYGN) SYNGENTA PARTICIPATIONS AG.  
PI Van Rooijen G, Helfetz PB, Briggs SP, Dalmia BK,  
PI Del Val G, Zapachinski S, Moloney M;  
XX P-PSDB; ABP60683.  
DR WPI; 2002-508806/54.  
XX  
XX Producing oil body associated with recombinant multimeric protein  
PT complex e.g. redox proteins and immunoglobulins comprises producing  
PT recombinant polypeptides capable of forming the complex in cells  
PT comprising oil bodies -  
XX  
XX Example 2: Page 169-171; 362pp; English.  
XX  
XX The present invention describes a method (M1) for producing an oil body  
CC associated with a recombinant multimeric protein complex (MPC). M1  
CC comprises producing in a cell comprising oil bodies a first and second  
CC recombinant polypeptide (P1, P2), where P1 is capable of associating  
CC with P2 to form the MPC and associating the complex with an occlusion  
CC body (OB) through an OB-targeting-protein capable of associating with OB  
CC and P1. M1 is useful for producing an oil body associated with a  
CC recombinant MPC. The oil bodies are further formulated for use in the  
CC preparation of a food product such as milk or wheat based food product,  
CC personal care product which reduces the oxidative stress on the surface  
CC area of the human body or used to lighten the skin, or a pharmaceutical  
CC composition used to treat chronic obstructive pulmonary disease (COPD),  
CC cataracts, diabetes, envenomation, bronchiopulmonary disease, psoriasis,  
CC malignancies, reperfusion injury, wound healing, sepsis, gastro  
CC intestinal (GI) bleeding, intestinal bowel disease (IBD), ulcers, GERD  
CC (gastro oesophageal reflux disease). ABN89569 to ABN89593 and ABP60677  
CC to ABP60964 represent sequence given in the exemplification of the  
CC present invention.  
XX  
XX Sequence 3888 BP; 1254 A; 722 C; 613 G; 1299 T; 0 other:  
SO  
Alignment Scores:  
Pred. No.: 5,84e-40 Length: 3888  
Score: 377.00 Matches: 73  
Percent Similarity: 71.67% Conservative: 13  
Best Local Similarity: 60.83% Mismatches: 34  
Query Match: 59.18% Indels: 0  
Gaps: 0  
DB: 24  
US-09-786-715-2 (1-122) x ABN89581 (1-3888)  
QY 2 AlaGluGluGlyGlnValIleAlaCysHisLysIleAspGluTrpGluGlyGlnLeuGly 21  
Db 1561 TCCGAAGAGAGCAAGATGCGCTGCCACCGTTGAGCATGGAACGACGAGCTTCAG 1620  
QY 22 LysTrpLysAspSerGluLysValIleValIleAspPheHisLysTrpGlyPro 41  
Db 1621 AAGGCTAATGATCCAAACTCTTGCGTGATTCAGGCTCTTGCTGCGACCA 1680  
QY 42 CysArgAlaIleAlaProTyrPheThrGluLeuAlaLysAsnAsnProAsnValAlaPhe 61  
Db 1681 TGTGCTTTCATCGCTCATCTTTGCTGATTTGGCTAAGAAACTTCCTACGCTTTC 1740  
QY 62 LeuLysValAlaAspGluLeuAsnSerValAlaSerLysTrpGluIleAsnAlaMet 81  
Db 1741 CTCGAAGTGTACTACTATGAAATGAAGTCGTCGCAAGTATGGGCGATGACGCGATG 1800  
QY 82 ProThrPheValPheLeuLysLysGlyLysIleIleGluLysIleValGlyAlaAspLys 101  
Db 1801 CCACCTTTCATGTTTGAAGGAGGAAAGATTGACAAAGTTGTTGAGCCAAAGAA 1860

QY 102 ValGlyLeuSerLysLysIleLeuGluLeuSerGlyThrThrProAlaAlaThrSerThr 121  
Db 1861 GATGAGCTTCACTCTACCATTTGCCAAACCTTGCTATGCGGATACAGCTAGAGCAACC 1920  
RESULT 5  
AAC37781  
ID AAC37781 standard; DNA: 470 BP.  
XX  
XX AAC37781;  
AC  
XX 17-OCR-2000 (first entry)  
DE Arabidopsis thaliana DNA fragment SEQ ID NO: 18644.  
XX  
XX Hybridisation assay; genetic mapping; gene expression control;  
KW protein identification; signal transduction pathway;  
XX metabolic pathway; promoter; termination sequence; ss.  
XX Arabidopsis thaliana.  
OS  
XX  
XX EP1033405-A2.  
PN  
XX  
PD 06-SEP-2000.  
XX  
XX 25-FEB-2000; 2000EP-0301439.  
XX  
XX 25-FEB-1999; 99US-0121825.  
PR 05-MAR-1999; 99US-0123180.  
PR 09-MAR-1999; 99US-0123548.  
PR 23-MAR-1999; 99US-0125788.  
PR 25-MAR-1999; 99US-0126264.  
PR 29-MAR-1999; 99US-0126785.  
PR 01-APR-1999; 99US-0127462.  
PR 06-APR-1999; 99US-0128234.  
PR 08-APR-1999; 99US-0128714.  
PR 16-APR-1999; 99US-0129845.  
PR 19-APR-1999; 99US-0130077.  
PR 21-APR-1999; 99US-0130449.  
PR 23-APR-1999; 99US-0130510.  
PR 23-APR-1999; 99US-0130891.  
PR 28-APR-1999; 99US-0131449.  
PR 30-APR-1999; 99US-0132048.  
PR 30-APR-1999; 99US-0132407.  
PR 04-MAY-1999; 99US-0132484.  
PR 05-MAY-1999; 99US-0132485.  
PR 06-MAY-1999; 99US-0132486.  
PR 07-MAY-1999; 99US-0132487.  
PR 07-MAY-1999; 99US-0132863.  
PR 11-MAY-1999; 99US-0134256.  
PR 14-MAY-1999; 99US-0134218.  
PR 14-MAY-1999; 99US-0134219.  
PR 14-MAY-1999; 99US-0134221.  
PR 14-MAY-1999; 99US-0134370.  
PR 18-MAY-1999; 99US-0134768.  
PR 19-MAY-1999; 99US-0134941.  
PR 20-MAY-1999; 99US-0135124.  
PR 21-MAY-1999; 99US-0135353.  
PR 24-MAY-1999; 99US-0135629.  
PR 25-MAY-1999; 99US-0136021.  
PR 27-MAY-1999; 99US-0136392.  
PR 28-MAY-1999; 99US-0136782.  
PR 01-JUN-1999; 99US-0137222.  
PR 03-JUN-1999; 99US-0137528.  
PR 04-JUN-1999; 99US-0137502.  
PR 07-JUN-1999; 99US-0137724.  
PR 08-JUN-1999; 99US-0138094.  
PR 10-JUN-1999; 99US-0138540.  
PR 10-JUN-1999; 99US-0138847.  
PR 14-JUN-1999; 99US-0139119.  
PR 16-JUN-1999; 99US-0139452.  
PR 17-JUN-1999; 99US-0139492.

PR 18-JUN-1999; 99US-0139454.  
PR 18-JUN-1999; 99US-0139455.  
PR 18-JUN-1999; 99US-0139456.  
PR 18-JUN-1999; 99US-0139457.  
PR 18-JUN-1999; 99US-0139458.  
PR 18-JUN-1999; 99US-0139459.  
PR 18-JUN-1999; 99US-0139460.  
PR 18-JUN-1999; 99US-0139461.  
PR 18-JUN-1999; 99US-0139462.  
PR 18-JUN-1999; 99US-0139463.  
PR 18-JUN-1999; 99US-0139750.  
PR 18-JUN-1999; 99US-0139763.  
PR 21-JUN-1999; 99US-0139817.  
PR 22-JUN-1999; 99US-0139899.  
PR 23-JUN-1999; 99US-0140353.  
PR 23-JUN-1999; 99US-0140354.  
PR 24-JUN-1999; 99US-0140695.  
PR 28-JUN-1999; 99US-0140823.  
PR 29-JUN-1999; 99US-0140991.  
PR 30-JUN-1999; 99US-0141287.  
PR 01-JUL-1999; 99US-0141842.  
PR 01-JUL-1999; 99US-0142154.  
PR 02-JUL-1999; 99US-0142055.  
PR 06-JUL-1999; 99US-0142390.  
PR 08-JUL-1999; 99US-0142803.  
PR 09-JUL-1999; 99US-0142920.  
PR 12-JUL-1999; 99US-0142977.  
PR 13-JUL-1999; 99US-0143542.  
PR 14-JUL-1999; 99US-0143624.  
PR 15-JUL-1999; 99US-0144005.  
PR 16-JUL-1999; 99US-0144085.  
PR 16-JUL-1999; 99US-0144086.  
PR 19-JUL-1999; 99US-0144325.  
PR 19-JUL-1999; 99US-0144332.  
PR 19-JUL-1999; 99US-0144333.  
PR 19-JUL-1999; 99US-0144334.  
PR 19-JUL-1999; 99US-0144335.  
PR 19-JUL-1999; 99US-0144352.  
PR 20-JUL-1999; 99US-0144632.  
PR 20-JUL-1999; 99US-0144884.  
PR 21-JUL-1999; 99US-0144814.  
PR 21-JUL-1999; 99US-0145086.  
PR 21-JUL-1999; 99US-0145088.  
PR 22-JUL-1999; 99US-0145085.  
PR 22-JUL-1999; 99US-0145087.  
PR 22-JUL-1999; 99US-0145089.  
PR 22-JUL-1999; 99US-0145192.  
PR 23-JUL-1999; 99US-0145145.  
PR 23-JUL-1999; 99US-0145218.  
PR 23-JUL-1999; 99US-0145224.  
PR 26-JUL-1999; 99US-0145276.  
PR 27-JUL-1999; 99US-0145919.  
PR 27-JUL-1999; 99US-0145918.  
PR 27-JUL-1999; 99US-0145919.  
PR 28-JUL-1999; 99US-0145951.  
PR 02-AUG-1999; 99US-0146386.  
PR 02-AUG-1999; 99US-0146388.  
PR 02-AUG-1999; 99US-0146389.  
PR 03-AUG-1999; 99US-0147038.  
PR 04-AUG-1999; 99US-0147204.  
PR 04-AUG-1999; 99US-0147302.  
PR 05-AUG-1999; 99US-0147192.  
PR 05-AUG-1999; 99US-0147260.  
PR 06-AUG-1999; 99US-0147303.  
PR 06-AUG-1999; 99US-0147416.  
PR 09-AUG-1999; 99US-0147493.  
PR 09-AUG-1999; 99US-0147935.  
PR 10-AUG-1999; 99US-0148171.  
PR 11-AUG-1999; 99US-0148319.  
PR 12-AUG-1999; 99US-0148341.  
PR 13-AUG-1999; 99US-0148565.  
PR 13-AUG-1999; 99US-0148684.

PR 16-AUG-1999; 99US-0149368.  
PR 17-AUG-1999; 99US-0149175.  
PR 18-AUG-1999; 99US-0149426.  
PR 20-AUG-1999; 99US-0149722.  
PR 20-AUG-1999; 99US-0149723.  
PR 20-AUG-1999; 99US-0149929.  
PR 23-AUG-1999; 99US-0149902.  
PR 23-AUG-1999; 99US-0149930.  
PR 25-AUG-1999; 99US-0150566.  
PR 26-AUG-1999; 99US-0150884.  
PR 27-AUG-1999; 99US-0151065.  
PR 27-AUG-1999; 99US-0151066.  
PR 30-AUG-1999; 99US-0151302.  
PR 31-AUG-1999; 99US-0151338.  
PR 01-SEP-1999; 99US-0151930.  
PR 07-SEP-1999; 99US-0152363.  
PR 10-SEP-1999; 99US-0153070.  
PR 13-SEP-1999; 99US-0153758.  
PR 15-SEP-1999; 99US-0154018.  
PR 16-SEP-1999; 99US-0154039.  
PR 20-SEP-1999; 99US-0154779.  
PR 22-SEP-1999; 99US-0155139.  
PR 23-SEP-1999; 99US-0155486.  
PR 24-SEP-1999; 99US-0155659.  
PR 28-SEP-1999; 99US-0156458.  
PR 29-SEP-1999; 99US-0156599.  
PR 04-OCT-1999; 99US-0157117.  
PR 05-OCT-1999; 99US-0157753.  
PR 06-OCT-1999; 99US-0157865.  
PR 07-OCT-1999; 99US-0158029.  
PR 08-OCT-1999; 99US-0158232.  
PR 12-OCT-1999; 99US-0158365.  
PR 13-OCT-1999; 99US-0158293.  
PR 13-OCT-1999; 99US-0159294.  
PR 13-OCT-1999; 99US-0159295.  
PR 14-OCT-1999; 99US-0159329.  
PR 14-OCT-1999; 99US-0159330.  
PR 14-OCT-1999; 99US-0159331.  
PR 14-OCT-1999; 99US-0159637.  
PR 14-OCT-1999; 99US-0159638.  
PR 18-OCT-1999; 99US-0159584.  
PR 21-OCT-1999; 99US-0160741.  
PR 21-OCT-1999; 99US-0160747.  
PR 21-OCT-1999; 99US-0160768.  
PR 21-OCT-1999; 99US-0160770.  
PR 21-OCT-1999; 99US-0160814.  
PR 22-OCT-1999; 99US-0160815.  
PR 22-OCT-1999; 99US-0160980.  
PR 22-OCT-1999; 99US-0160981.  
PR 22-OCT-1999; 99US-0160989.  
PR 25-OCT-1999; 99US-0161404.  
PR 25-OCT-1999; 99US-0161405.  
PR 25-OCT-1999; 99US-0161406.  
PR 26-OCT-1999; 99US-0161359.  
PR 26-OCT-1999; 99US-0161360.  
PR 26-OCT-1999; 99US-0161361.  
PR 28-OCT-1999; 99US-0161920.  
PR 28-OCT-1999; 99US-0161992.  
PR 28-OCT-1999; 99US-0161993.  
PR 29-OCT-1999; 99US-0162142.

## Alignment Scores:

Pred. No.: 3.3e-41  
Score: 376.50  
Percent Similarity: 73.33%  
Best Local Similarity: 60.83%  
Query Match: 59.11%  
DB: 21  
Length: 470  
Matches: 73  
Conservative: 15  
Mismatches: 25  
Indels: 7  
Gaps: 1

US-09-786-715-2 (1-122) x AAC37781 (1-470)

OY 2 AlaglugluglInvalItleAlacyshtstslsyleaspglutprgluglInleucly 21

Db 101 TCGAAGAAAGACAGATCGCTGCCACACCGTTGACATGGAACGACGCTTCAG 160  
QY 22 LysTrpLysAspSerGluLysLeuValValAlaAspPheThrAlaSerTrpCysGlyPro 41  
Db 161 AAGGCTAATGAAATCCAAACTCTTGCTGCTGATTCACGGCTTCTTGCTGACCA 220  
QY 42 CysArgAlaIleAlaProTyrPheThrGluLeuAlaLysAsnAsnProAsnValAlaPhe 61  
Db 221 TGTGCTTCATGCTCATCTTCTTGCTGATTCGCTAAGAAACCTTCTTACGCTTTC 280  
QY 62 LeuLysValAspValAspGluLeuAsnSerValAlaSerLysTrpGluIleAsnAlaMet 81  
Db 281 CTCGAAGCTTGATCTGATGATGAATGAAGTCGTCGCAAGTATGCGGATACAGCGCATG 340  
QY 82 ProThrPheValPheLeuLysLysGlyLysIleIleGluLysIleValAlaAspLys 101  
Db 341 CCACCTTCATGTTTGGAAAGAAAGGAGATTTTGGACAAACTGTTGGAGCCACAGAAA 400  
QY 102 ValGlyLeuSerLys-----LysIleLeuGluLeuSerGlyThr 114  
Db 401 GATGACCTTCAGTCTACCATTTGCCAAACACTTGCCTCGCTATTACCAAGTTTAAAGAAC 460  
RESULT 6  
AB212358 standard; DNA: 345 BP.  
ID AB212358  
AC AB212358;  
XX 21-JUN-2003 (first entry)  
DE Arabidopsis thaliana stress regulated gene SEQ ID NO 163.  
XX Arabidopsis thaliana; plant; gene; stress; transgenic; ds.  
OS Arabidopsis thaliana.  
XX WO200216655-A2.  
PN 28-FEB-2002.  
PD 24-AUG-2001; 2001MO-US26685.  
PF 24-AUG-2000; 2000US-227866P.  
XX 26-JUN-2001; 2001US-264647P.  
PR 22-JUN-2001; 2001US-300111P.  
XX (SCRI ) SCRIPPS RES INST.  
PA (SYGN ) SYNGENTA PARTICIPATIONS AG.  
XX Harper JF, Kreps J, Wang X, Zhu T;  
PI WPI; 2002-304127/34.  
DR Identifying a stress condition to which a plant cell has been exposed  
XX and producing plants with increased tolerance to these abiotic stresses  
PT -  
PI Claim 144; SEQ ID NO 163; 577pp + Sequence Listing; English.  
PS The invention relates to identifying a stress condition to which a plant  
XX cell has been exposed, comprising:  
CC (a) contacting nucleic acid representative of expressed polynucleotides  
CC in the plant cell with an array or probes representative of the plant  
CC cell genome; and  
CC (b) detecting a profile of expressed polynucleotides in the plant cell  
CC characteristic of a stress response. The method is useful in the  
CC production of transgenic plants, cells and seeds and in producing plants  
CC with increased tolerance to abiotic stress. The present sequence is that  
CC of an Arabidopsis thaliana stress regulated gene (AB212196-AB217574) used  
CC in methods of the invention.  
CC Note: The sequence data for this patent is not represented in the printed  
CC specification but is based on sequence information supplied to Derwent by

CC the European Patent Office.  
XX Sequence 345 BP; 90 A; 68 C; 90 G; 97 T; 0 other;  
SQ Alignment Scores:  
Pred. No.: 2.47e-41 Length: 345  
Score: 376.00 Matches: 71  
Percent Similarity: 77.57% Conservative: 12  
Best Local Similarity: 66.36% Mismatches: 24  
Query Match: 59.03% Indels: 0  
DB: 24 Gaps: 0  
US-09-786-715-2 (1-122) x AB212358 (1-345)  
QY 2 AlagluGluLysIleValIleAlaCysHisLysIleAspGluTrpGluGluIleGly 21  
Db 7 TCGAAGAAAGACAGATCGCTGCCACACCGTTGACATGGAACGACGCTTCAG 66  
QY 22 LysTrpLysAspSerGluLysLeuValValAlaAspPheThrAlaSerTrpCysGlyPro 41  
Db 67 AAGGCTAATGAAATCCAAACTCTTGCTGCTGATTCACGGCTTCTTGCTGACCA 126  
QY 42 CysArgAlaIleAlaProTyrPheThrGluLeuAlaLysAsnAsnProAsnValAlaPhe 61  
Db 127 TGTGCTTCATGCTCATCTTCTTGCTGATTCGCTAAGAAACCTTCTTACGCTTTC 186  
QY 62 LeuLysValAspValAspGluLeuAsnSerValAlaSerLysTrpGluIleAsnAlaMet 81  
Db 187 CTCGAAGCTTGATCTGATGATGAATGAAGTCGTCGCAAGTATGCGGATACAGCGCATG 246  
QY 82 ProThrPheValPheLeuLysLysGlyLysIleIleGluLysIleValAlaAspLys 101  
Db 247 CCACCTTCATGTTTGGAAAGAAAGGAGATTTTGGACAAACTGTTGGAGCCACAGAAA 306  
QY 102 ValGlyLeuSerLysLysIle 108  
Db 307 GATGACCTTCAGTCTACCATTT 327  
RESULT 7  
ABN89587 standard; DNA: 345 BP.  
ID ABN89587  
AC ABN89587;  
XX 06-SEP-2002 (first entry)  
DE Arabidopsis thaliana thioredoxin h (Trx h 1) DNA SEQ ID NO:38.  
XX Multimeric protein; redox protein; thioredoxin; thioredoxin reductase;  
KW oil body; ophthalmological; antidiabetic; cytoskeletal; antiproliferic;  
KW vasotropic; vulnary; antibacterial; immunosuppressive; antitumor;  
KW food product; milk; wheat; oxidative stress; cataract; diabetes;  
KW chronic obstructive pulmonary disease; envenomation; psoriasis; sepsis;  
KW bronchiopulmonary disease; malignancy; reperfusion injury; wound healing;  
KW gastro intestinal bleeding; intestinal bowel disease; ulcer;  
KW gastro oesophageal reflux disease; gene; ds.  
XX Arabidopsis thaliana.  
OS WO200250289-A1.  
PN 27-JUN-2002.  
PD 19-DEC-2001; 2001MO-US50240.  
PF 19-DEC-2000; 2000US-0742900.  
XX 04-JUL-2001; 2001US-302885P.  
PR 04-DEC-2001; 2001US-0006038.  
XX (SEMB-) SEMBIOSYS GENETICS INC.  
PA (SYGN ) SYNGENTA PARTICIPATIONS AG.  
XX Van Rooijen G, Deckers H, Helfetz PB, Briggs SP, Dalmia BK;  
PI





PR 18-JUN-1999; 99US-0139763.  
PR 21-JUN-1999; 99US-0139817.  
PR 22-JUN-1999; 99US-0139899.  
PR 23-JUN-1999; 99US-0140353.  
PR 23-JUN-1999; 99US-0140354.  
PR 24-JUN-1999; 99US-0140695.  
PR 28-JUN-1999; 99US-0140823.  
PR 29-JUN-1999; 99US-0140991.  
PR 30-JUN-1999; 99US-0141287.  
PR 01-JUL-1999; 99US-0141842.  
PR 01-JUL-1999; 99US-0142154.  
PR 02-JUL-1999; 99US-0142055.  
PR 06-JUL-1999; 99US-0142390.  
PR 08-JUL-1999; 99US-0142803.  
PR 09-JUL-1999; 99US-0142920.  
PR 12-JUL-1999; 99US-0142977.  
PR 13-JUL-1999; 99US-0143542.  
PR 14-JUL-1999; 99US-0143644.  
PR 15-JUL-1999; 99US-0144005.  
PR 16-JUL-1999; 99US-0144085.  
PR 16-JUL-1999; 99US-0144086.  
PR 19-JUL-1999; 99US-0144325.  
PR 19-JUL-1999; 99US-0144331.  
PR 19-JUL-1999; 99US-0144332.  
PR 19-JUL-1999; 99US-0144333.  
PR 19-JUL-1999; 99US-0144334.  
PR 19-JUL-1999; 99US-0144335.  
PR 20-JUL-1999; 99US-0144352.  
PR 20-JUL-1999; 99US-0144632.  
PR 20-JUL-1999; 99US-0144684.  
PR 21-JUL-1999; 99US-0144814.  
PR 21-JUL-1999; 99US-0145086.  
PR 21-JUL-1999; 99US-0145088.  
PR 22-JUL-1999; 99US-0145085.  
PR 22-JUL-1999; 99US-0145087.  
PR 22-JUL-1999; 99US-0145089.  
PR 22-JUL-1999; 99US-0145192.  
PR 23-JUL-1999; 99US-0145145.  
PR 23-JUL-1999; 99US-0145218.  
PR 23-JUL-1999; 99US-0145224.  
PR 26-JUL-1999; 99US-0145276.  
PR 27-JUL-1999; 99US-0145913.  
PR 27-JUL-1999; 99US-0145918.  
PR 27-JUL-1999; 99US-0145919.  
PR 28-JUL-1999; 99US-0145951.  
PR 02-AUG-1999; 99US-0146386.  
PR 02-AUG-1999; 99US-0146388.  
PR 03-AUG-1999; 99US-0147038.  
PR 04-AUG-1999; 99US-0147204.  
PR 04-AUG-1999; 99US-0147302.  
PR 05-AUG-1999; 99US-0147192.  
PR 05-AUG-1999; 99US-0147260.  
PR 06-AUG-1999; 99US-0147303.  
PR 06-AUG-1999; 99US-0147416.  
PR 09-AUG-1999; 99US-0147493.  
PR 09-AUG-1999; 99US-0147935.  
PR 10-AUG-1999; 99US-0148171.  
PR 11-AUG-1999; 99US-0148319.  
PR 12-AUG-1999; 99US-0148341.  
PR 13-AUG-1999; 99US-0148565.  
PR 13-AUG-1999; 99US-0148684.  
PR 16-AUG-1999; 99US-0149368.  
PR 17-AUG-1999; 99US-0149175.  
PR 18-AUG-1999; 99US-0149126.  
PR 20-AUG-1999; 99US-0149722.  
PR 20-AUG-1999; 99US-0149723.  
PR 20-AUG-1999; 99US-0149929.  
PR 23-AUG-1999; 99US-0149902.  
PR 23-AUG-1999; 99US-0149930.  
PR 25-AUG-1999; 99US-0150566.  
PR 26-AUG-1999; 99US-0150884.  
PR 27-AUG-1999; 99US-0151065.

PR 27-AUG-1999; 99US-0151066.  
PR 27-AUG-1999; 99US-0151080.  
PR 30-AUG-1999; 99US-0151303.  
PR 31-AUG-1999; 99US-0151438.  
PR 01-SEP-1999; 99US-0151930.  
PR 07-SEP-1999; 99US-0151930.  
PR 10-SEP-1999; 99US-0152363.  
PR 13-SEP-1999; 99US-0153758.  
PR 15-SEP-1999; 99US-0154018.  
PR 16-SEP-1999; 99US-0154039.  
PR 20-SEP-1999; 99US-0154779.  
PR 22-SEP-1999; 99US-0155139.  
PR 23-SEP-1999; 99US-0155486.  
PR 24-SEP-1999; 99US-0155659.  
PR 28-SEP-1999; 99US-0156458.  
PR 29-SEP-1999; 99US-0156596.  
PR 04-OCT-1999; 99US-0157117.  
PR 05-OCT-1999; 99US-0157753.  
PR 06-OCT-1999; 99US-0157865.  
PR 07-OCT-1999; 99US-0158029.  
PR 08-OCT-1999; 99US-0158232.  
PR 12-OCT-1999; 99US-0158369.  
PR 13-OCT-1999; 99US-0159293.  
PR 13-OCT-1999; 99US-0159294.  
PR 13-OCT-1999; 99US-0159295.  
PR 14-OCT-1999; 99US-0159329.  
PR 14-OCT-1999; 99US-0159330.  
PR 14-OCT-1999; 99US-0159331.  
PR 14-OCT-1999; 99US-0159637.  
PR 14-OCT-1999; 99US-0159638.  
PR 18-OCT-1999; 99US-0159684.  
PR 21-OCT-1999; 99US-0160741.  
PR 21-OCT-1999; 99US-0160767.  
PR 21-OCT-1999; 99US-0160768.  
PR 21-OCT-1999; 99US-0160770.  
PR 21-OCT-1999; 99US-0160814.  
PR 21-OCT-1999; 99US-0160815.  
PR 22-OCT-1999; 99US-0160980.  
PR 22-OCT-1999; 99US-0160981.  
PR 22-OCT-1999; 99US-0160989.  
PR 25-OCT-1999; 99US-0161404.  
PR 25-OCT-1999; 99US-0161405.  
PR 25-OCT-1999; 99US-0161406.  
PR 26-OCT-1999; 99US-0161359.  
PR 26-OCT-1999; 99US-0161360.  
PR 26-OCT-1999; 99US-0161361.  
PR 28-OCT-1999; 99US-0161920.  
PR 28-OCT-1999; 99US-0161992.  
PR 28-OCT-1999; 99US-0161993.  
PR 29-OCT-1999; 99US-0162142.

## Alignment Scores:

Pred. No.:	4,96e-41	Length:	561
Score:	376.00	Matches:	71
Percent Similarity:	77.57%	Conservative:	12
Best Local Similarity:	66.36%	Mismatches:	24
Query Match:	59.03%	Indels:	0
DB:	21	Gaps:	0

US-09-786-715-2 (1-122) x AACS1522 (1-561)

QY 2 AAlaGlUGlUGlGlnValIleAlaCysHisLysIleAspGluTrpGluGlnLeuGly 21  
Db 100 TCGAGAGAAAGCAAGATCGCTGCCACACCGTTGACATGGAAGCAAGCACTTCAG 159  
QY 22 LysTrpLysAspSerGluLysLeuValValAspPheThrAsaSerTrpCysGlyPro 41  
Db 160 AAGGCTAATGAATCAAAACCTGTGGTGTTGATTTCACGCTCTTGTGTGACCA 219  
QY 42 CysArgAlaIleLeaProTyRpheThrGluLeuAlaLysAsnAsnProAsnValAlaPhe 61  
Db 220 TGTGTTTCATCGCTCCATTCCTTTGCGATTTGGCTAAGAACTTCTTAACGTCCTTTC 279

[illegible]

```

PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148664.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-015659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160860.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

```

Alignment Scores:  
Pred. No.:  
Score:  
Percent Similarity:

4.99e-41  
376.00  
77.578

Length: 563  
Matches: 71  
Conservative: 12

```

Best Local Similarity: 66.36% Mismatches: 24
Query Match: 59.03% Indels: 0
DB: 21 Gaps: 0
US-09-786-715-2 (1-122) x AAC34121 (1-563)

QY 2 AlaGluGluGlnValIleAlaCysHisLysIleAspGluTrpGluGlnLeuGly 21
   ::::::::::::::::::::::::::::::::::::
Db 102 TCGAAGACAGACAAAGATCGCGCCACACCGTTGAGACATGACAGACAGACCTTCAG 161
   ::::::::::::::::::::::::::::::::::::::::::::

QY 22 LysTrpLysAspSerGluLysLeuValValAlaAspPheThrAlaSerTrpCysGlyPro 41
   ::::::::::::::::::::::::::::::::::::::::::::::::::::
Db 162 AAGCCTAATGATCACTTGTGATGATGATGATGATGATGATGATGATGATGATGATG 221
   ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::

QY 42 CysArgAlaIleAlaProTyrPheThrGluLeuAlaLysAsnAsnProAsnValAlaPhe 61
   ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
Db 222 TGTCGTTTCATCGCTCATCTTTCGATTTGGCTTAAGAAACTCTTAAGCTCTTTC 281
   ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::

QY 62 LeuLysValAspValAspGluLeuAsnSerValAlaSerLysTrpGluIleAsnAlaMet 81
   ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
Db 282 CTCAGAGCTGATACTGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 341

QY 82 ProThrPheValPheLeuLysGlyLysIleIleGluLysIleValGlyAlaAspLys 101
   ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
Db 342 CCACCTTCATGTTTGAAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 401

QY 102 ValGlyLeuSerLysIle 108
   ::::
Db 402 GATGAGCTTCAGTCTACCTT 422

RESULT 10
ABBS3095
ID ABBS3095 standard; DNA; 3129 BP.
XX
AC ABBS3095;
XX
DT 29-NOV-2002 (first entry)
XX
XX DNA encoding thioredoxin-phaseolin fusion protein.
XX
XX Thioredoxin; thioredoxin reductase; gene expression; oleosin;
XX oil body; gene; ds; thioredoxin reductase; phaseolin.
XX
XX Arabidopsis thaliana.
XX
XX Phaseolus vulgaris.
XX
XX Synthetic.
XX
FH Key Location/Qualifiers
FT CDS 1555..1899
/*tag= a
/*product= "thioredoxin-phaseolin fusion protein"

US2002088025-A1.
04-JUL-2002.
XX
XX
XX 03-JUL-2001; 2001US-0897425.
XX
XX
XX 22-FEB-1991; 91US-0659835.
XX 16-NOV-1993; 93US-0142418.
XX 30-DEC-1994; 94US-0366783.
XX 25-APR-1997; 97US-0846021.
XX 18-DEC-1998; 98US-0210843.
XX
XX (MOLONEY) MOLONEY M. M.
XX (DALMIA) DALMIA B. K.
XX
XX Moloney MM, Dalmlia BK;
XX
XX WPI: 2002-635723/68.
XX P-PSDB; ABG32916.
XX
XX Expressing protein, by introducing chimeric nucleotide regulatory

```

PT sequence, sequence encoding fusion protein, having sequence encoding  
PT protein, oleosin gene and sequence encoding termination region and  
PT producing protein

XX Example 21; Fig 12; 69pp; English.

XX The invention describes a method of expressing thioredoxin or thioredoxin  
CC reductase (I) in the oil body of a host cell using an oil body protein  
CC gene. The method involves introducing a chimeric nucleic acid comprising  
CC a first sequence to regulate transcription, a second DNA sequence  
CC encoding a fusion polypeptide, comprising a sequence encoding an oleosin  
CC gene and sequence encoding (I) and a third sequence encoding a  
CC termination region functional in the host cell and growing the host cell  
CC to produce a fusion polypeptide. The method or (I) is useful for  
CC expression of a thioredoxin or thioredoxin reductase by a host cell. This  
CC sequence represents an Arabidopsis thaliana thioredoxin gene controlled  
CC by a phaseolin promoter and terminator for use in an expression vector.

XX Sequence 3129 BP; 1051 A; 549 C; 439 G; 1090 T; 0 other;

Alignment Scores:

Pred. No.:	5,84e-40	Length:	3129
Score:	376.00	Matches:	71
Percent Similarity:	77.57%	Conservative:	12
Best Local Similarity:	66.36%	Mismatches:	24
Query Match:	59.03%	Indels:	0
DB:	24	Gaps:	0

US-09-786-715-2 (1-122) x AB853095 (1-3129)

OY 2 AAGAGUGUGIYGVNValIleAlaCysHISLysIleAspGLUTPrpLugLynLeugly 21

DB 1561 TCGGAAGAGACAACTGATCGCTGCCACACCGTTGACACATGACAGCAGCTTCAG 1620

OY 22 LysTrpLysAspSerGluLysLeuValValAlaSpPheThrAlaSerTrpCysGlyPro 41

DB 1621 AAGGTAATGATCCAAACTCTTGTGGTGGATTTCACGGCTCTTGTGGTGGACCA 1680

OY 42 CysArgAlaIleAlaPheTrpThrGluLeuAlaLysAsnProAsnValAlaPhe 61

DB 1681 TGTGCTTCATCGCTCCATCTTGTGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1740

OY 62 LeuLysValAspValAspLysLeuAsnSerValAlaSerLysTrpLysLeuAlaMet 81

DB 1741 CTCAGGTTGATACATGATGATGAGAGTGGCGCAGATGATGGCGCATACAGCGCATG 1800

OY 82 ProThrPheValPheLeuLysLysGlyLysIleIleGluLysIleValAlaLysAspLys 101

DB 1801 CCAACCTTCATGTTTGAAGAGAGGAGATTGTCACAAAGTTGTGGAGCCAGAGAA 1860

OY 102 ValGlyLeuSerLysLysIle 108

DB 1861 GATGAGCTTCAGCTACCAATT 1881

DB 1861 GATGAGCTTCAGCTACCAATT 1881

DB 1861 GATGAGCTTCAGCTACCAATT 1881

DB 1861 GATGAGCTTCAGCTACCAATT 1881

DB 1861 GATGAGCTTCAGCTACCAATT 1881

DB 1861 GATGAGCTTCAGCTACCAATT 1881

DB 1861 GATGAGCTTCAGCTACCAATT 1881

DB 1861 GATGAGCTTCAGCTACCAATT 1881

DB 1861 GATGAGCTTCAGCTACCAATT 1881

DB 1861 GATGAGCTTCAGCTACCAATT 1881

OS Arabidopsis sp.  
OS Phaseolus vulgaris.

XX WO200250289-A1.

XX 27-JUN-2002.

XX 19-DEC-2001; 2001MO-US50240.

XX 19-DEC-2000; 2000US-0742900.

XX 05-JUL-2001; 2001US-302885P.

XX 04-DEC-2001; 2001US-0006038.

XX (SEMB-) SEMBIOSYS GENETICS INC.

XX (SYGN) SYNGENTA PARTICIPATIONS AG.

XX Van Rooijen G, Deckers H, Helfetz PB, Briggs SP, Dalmia BK;

XX Del Val G, Zaplachinski S, Moloney M;

XX WPI: 2002-508806/54.

XX P-PSDB; AB60680.

XX Producing oil body associated with recombinant multimeric protein

XX complex e.g. redox proteins and immunoglobulins comprises producing

XX recombinant polypeptides capable of forming the complex in cells

XX comprising oil bodies -

XX Example 2; Page 165-166; 362pp; English.

XX The present invention describes a method (M1) for producing an oil body

XX associated with a recombinant multimeric protein complex (MPC). M1

XX comprises producing in a cell comprising oil bodies a first and second

XX recombinant polypeptide (P1, P2), where P1 is capable of associating

XX with P2 to form the MPC and associating the complex with an occlusion

XX body (OB) through an OB-targeting-protein capable of associating with OB

XX and P1. M1 is useful for producing an oil body associated with a

XX recombinant MPC. The oil bodies are further formulated for use in the

XX preparation of a food product such as milk or wheat based food product,

XX personal care product which reduces the oxidative stress on the surface

XX area of the human body or used to lighten the skin, or a pharmaceutical

XX composition used to treat chronic obstructive pulmonary disease (COPD),

XX cataracts, diabetes, emphysema, bronchiolopulmonary disease, psoriasis,

XX malnancies, reperfusion injury, wound healing, sepsis, gastro

XX intestinal (GI) bleeding, intestinal bowel disease (IBD), ulcers, GERD

XX (gastro oesophageal reflux disease). AB89569 to AB89593 and AB860677

XX to AB860964 represent sequence given in the exemplification of the

XX present invention.

XX Sequence 3129 BP; 1051 A; 549 C; 439 G; 1090 T; 0 other;

XX Alignment Scores:

Pred. No.:	5,84e-40	Length:	3129
Score:	376.00	Matches:	71
Percent Similarity:	77.57%	Conservative:	12
Best Local Similarity:	66.36%	Mismatches:	24
Query Match:	59.03%	Indels:	0
DB:	24	Gaps:	0

US-09-786-715-2 (1-122) x AB89579 (1-3129)

OY 2 AAGAGUGUGIYGVNValIleAlaCysHISLysIleAspGLUTPrpLugLynLeugly 21

DB 1561 TCGGAAGAGACAACTGATCGCTGCCACACCGTTGACACATGACAGCAGCTTCAG 1620

OY 22 LysTrpLysAspSerGluLysLeuValValAlaSpPheThrAlaSerTrpCysGlyPro 41

DB 1621 AAGGTAATGATCCAAACTCTTGTGGTGGATTTCACGGCTCTTGTGGTGGACCA 1680

OY 42 CysArgAlaIleAlaPheTrpThrGluLeuAlaLysAsnProAsnValAlaPhe 61

DB 1681 TGTGCTTCATCGCTCCATCTTGTGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1740

OY 62 LeuLysValAspValAspLysLeuAsnSerValAlaSerLysTrpLysLeuAlaMet 81



```

PR 19-DEC-2000; 2000US-0742900.
PR 05-JUL-2001; 2001US-302885P.
PR 04-DEC-2001; 2001US-0006038.
XX
XX (SEMB-) SEMBIOSYS GENETICS INC.
XX (SYGN) SYNGENTA PARTICIPATIONS AG.
PI Van Rooijen G, Deckers H, Helfetz PB, Briggs SP, Dalmia BK;
PI Del Val G, Zaplachinski S, Moloney M;
XX
XX WPI: 2002-508806/54.
DR P-PSDB: ABP60677.
XX
XX Producing oil body associated with recombinant multimeric protein
XX complex e.g. redox proteins and immunoglobulins comprises producing
XX recombinant polypeptides capable of forming the complex in cells
XX comprising oil bodies -
XX
XX Example 2; Page 167-168; 362pp; English.
XX
XX The present invention describes a method (M1) for producing an oil body
XX associated with a recombinant multimeric protein complex (MPC). M1
XX comprises producing in a cell comprising oil bodies a first and second
XX recombinant polypeptide (P1, P2), where P1 is capable of associating
XX with P2 to form the MPC and associating the complex with an occlusion
XX body (OB) through an OB-targeting-protein capable of associating with OB
XX and P1. M1 is useful for producing an oil body associated with a
XX recombinant MPC. The oil bodies are further formulated for use in the
XX preparation of a food product such as milk or wheat based food product,
XX personal care product which reduces the oxidative stress on the surface
XX area of the human body or used to lighten the skin, or a pharmaceutical
XX composition used to treat chronic obstructive pulmonary disease (COPD),
XX cataracts, diabetes, envenomation, bronchiopulmonary disease, psoriasis,
XX malignancies, reperfusion injury, wound healing, sepsis, gastro
XX intestinal (GI) bleeding, intestinal bowel disease (IBD), ulcers, GERD
XX (gastro oesophageal reflux disease). ABN89569 to ABN89593 and ABP60677
XX to ABP60964 represent sequence given in the exemplification of the
XX present invention.
XX
XX Sequence 3888 BP; 1254 A; 723 C; 613 G; 1298 T; 0 other;
SQ
Alignment Scores:
Pred. No.: 7.97e-40 Length: 3888
Score: 376.00 Matches: 71
Percent Similarity: 77.57% Conservative: 12
Best Local Similarity: 66.36% Mismatches: 24
Query Match: 59.03% Indels: 0
DB: 24 Gaps: 0
US-09-786-715-2 (1-122) x ABN89580 (1-3888)
OY 2 Alaglugluglgnlvalilialacyshtslslleapglutrpdlnglgnleucl 21
DB 2320 TCGGAGGAGGAGCAAGTATGCGCCGACCCGTTGAGACATGGACAGCAGCTTAG 2379
OY 22 Lystriplyaspserglulslleuvalvalalasphefnrlalasertrpcysglypro 41
DB 2380 AAGGTATATGATCAACCAACCTTGTGCTGATTTGCTACAGCGCTTGTGGTGGACCA 2439
OY 42 CysArgAlaIleAlaProTyrPheThrGluLeuAlaLysAsnAsnProAsnValAlaIle 61
DB 2440 TGTGCTTCATCGCCATCTTGTGCTGATTTGGCTAGAAACCTTCTTGTGGTGGTTC 2499
OY 62 LeuylsValaspValaspGluLeuAsnserValAlaSerLysTrpGluIleAsnAlaMet 81
DB 2500 CTCAGGTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 2559
OY 82 ProThrPheValAlaPheLeuLysGlyLysLysLysLysLysLysLysLysLysLys 101
DB 2560 CCACCTTCATGCTTTTTCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 2619
OY 102 ValGlyLeuSerLysLysLysLysLysLysLysLysLysLysLysLysLysLysLys 108

```

```

DB 2620 GATGAGCTTCACTTACCATT 2640
RESULT 14
ABN89586
XX ABN89586 standard; DNA; 4935 BP;
XX
XX ABN89586;
XX
XX 06-SEP-2002 (first entry)
XX
XX Promoter-Oleoin thioedoxin reductase-linker-thioedoxin-terminator #33.
XX
XX Multimeric protein; redox protein; thioedoxin; thioedoxin reductase;
XX oil body; ophthalmological; antidiabetic; cytosolic; antiproliferative;
XX vasotropic; vulnary; antibacterial; immunosuppressive; antileuk;
XX food product; milk; wheat; oxidative stress; cataract; diabetes;
XX chronic obstructive pulmonary disease; envenomation; psoriasis; sepsis;
XX bronchiopulmonary disease; malignancy; reperfusion injury; wound healing;
XX gastro intestinal bleeding; intestinal bowel disease; ulcer;
XX gastro oesophageal reflux disease; gene; ds.
XX
XX Arabidopsis sp.
XX Phaseolus vulgaris.
XX
XX WO200250289-A1.
XX
XX 27-JUN-2002.
XX
XX 19-DEC-2001; 2001WO-US50240.
XX
XX 19-DEC-2000; 2000US-0742900.
XX 05-JUL-2001; 2001US-302885P.
XX 04-DEC-2001; 2001US-0006038.
XX
XX (SEMB-) SEMBIOSYS GENETICS INC.
XX (SYGN) SYNGENTA PARTICIPATIONS AG.
XX
XX Van Rooijen G, Deckers H, Helfetz PB, Briggs SP, Dalmia BK;
XX Del Val G, Zaplachinski S, Moloney M;
XX
XX WPI: 2002-508806/54.
XX P-PSDB: ABP60692.
XX
XX Producing oil body associated with recombinant multimeric protein
XX complex e.g. redox proteins and immunoglobulins comprises producing
XX recombinant polypeptides capable of forming the complex in cells
XX comprising oil bodies -
XX
XX Example 2; Page 189-192; 362pp; English.
XX
XX The present invention describes a method (M1) for producing an oil body
XX associated with a recombinant multimeric protein complex (MPC). M1
XX comprises producing in a cell comprising oil bodies a first and second
XX recombinant polypeptide (P1, P2), where P1 is capable of associating
XX with P2 to form the MPC and associating the complex with an occlusion
XX body (OB) through an OB-targeting-protein capable of associating with OB
XX and P1. M1 is useful for producing an oil body associated with a
XX recombinant MPC. The oil bodies are further formulated for use in the
XX preparation of a food product such as milk or wheat based food product,
XX personal care product which reduces the oxidative stress on the surface
XX area of the human body or used to lighten the skin, or a pharmaceutical
XX composition used to treat chronic obstructive pulmonary disease (COPD),
XX cataracts, diabetes, envenomation, bronchiopulmonary disease, psoriasis,
XX malignancies, reperfusion injury, wound healing, sepsis, gastro
XX intestinal (GI) bleeding, intestinal bowel disease (IBD), ulcers, GERD
XX (gastro oesophageal reflux disease). ABN89569 to ABN89593 and ABP60677
XX to ABP60964 represent sequence given in the exemplification of the
XX present invention.
XX
XX Sequence 4935 BP; 1528 A; 932 C; 898 G; 1577 T; 0 other;
SQ
Alignment Scores:
Pred. No.: 1.53e-39 Length: 4935

```





GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

## OM protein - nucleic search, using frame\_plus\_p2n model

Run on: August 17, 2003, 19:18:57 ; Search time 45.2132 Seconds  
(without alignments)  
1190.995 Million cell updates/sec

Title: US-09-786-715-2

Perfect score: 637

Sequence: 1 MAERGVIACHRIDWEGOL.....GLSKRIELSGTTPATSTA 122

Scoring table: BLOSUM62  
Xgapop 10.0, Xgapext 0.5  
Ygapop 10.0, Ygapext 0.5  
Fgapop 6.0, Fgapext 7.0  
Delop 6.0, Delext 7.0

Searched: 569978 seqs, 220691566 residues

Total number of hits satisfying chosen parameters: 1139956

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 08  
Maximum Match 1008  
Listing first 45 summaries

## Command line parameters:

-MODEL=frame+P2n.model -DEV=xlp  
-O=/cgn2.1/USPTO\_spool/US09786715/runat\_11082003.150515.6083/app.query.fasta\_1.1052  
-DB-issued.Patents.NA -QEMT=fastap -SUFFIX=rii -MINMATCH=0.1 -LOOPCL=0  
-LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=bl0sum62 -TRANS=human40.cdi  
-LIST=45 -DOCALLIGN=200 -THR.SCORE=pct -THR.MAX=100 -THR.MIN=0 -ALIGN=15  
-MODE=LOCAL -OUTFMT=ptio -NORR=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000  
-USER=US09786715.ecgn.1.1.175-@runat\_11082003.150515.6083 -NCPU=6 -ICPU=3  
-NO\_MMAP -LARGEOUTERY -NEG.SCORES=0 -WAIT -DSPELOCK=100 -LONGLOC  
-DEV.TIMEOUT=120 -WARN.TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -Fgapop=6  
-Fgapext=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

## Database :

- Issued Patents.NA:\*
- 1: /cgn2.6/ptodata/1/ina/5A.COMB.seq:\*
  - 2: /cgn2.6/ptodata/1/ina/5B.COMB.seq:\*
  - 3: /cgn2.6/ptodata/1/ina/6A.COMB.seq:\*
  - 4: /cgn2.6/ptodata/1/ina/6B.COMB.seq:\*
  - 5: /cgn2.6/ptodata/1/ina/PCtus.COMB.seq:\*
  - 6: /cgn2.6/ptodata/1/ina/backfiles1.seq:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match Length	ID	Description
1	335.5	52.7	653 1 US-08-181-271A-103	Sequence 103, App
2	335.5	52.7	653 1 US-08-449-315-103	Sequence 103, App
3	335.5	52.7	653 1 US-08-444-803-103	Sequence 103, App
4	335.5	52.7	653 1 US-08-449-043-103	Sequence 103, App
5	335.5	52.7	653 1 US-08-456-265A-103	Sequence 103, App
6	335.5	52.7	653 1 US-08-455-416-103	Sequence 103, App
7	335.5	52.7	653 1 US-08-455-244-103	Sequence 103, App
8	335.5	52.7	653 1 US-08-454-876-103	Sequence 103, App
9	335.5	52.7	653 2 US-08-457-364-103	Sequence 103, App
10	335.5	52.7	653 2 US-08-456-262-103	Sequence 103, App
11	335.5	52.7	653 2 US-08-456-240-103	Sequence 103, App
12	335.5	52.7	653 2 US-08-455-736-103	Sequence 103, App

13	335.5	52.7	653 2 US-08-971-217-103	Sequence 103, App
14	335.5	52.7	653 3 US-09-350-600-103	Sequence 103, App
15	302	47.4	369 4 US-09-540-014-1	Sequence 1, Appl1
16	298	46.8	382 4 US-09-540-014-3	Sequence 3, Appl1
17	298	46.8	393 4 US-09-540-014-5	Sequence 5, Appl1
18	231	36.3	557 4 US-09-404-879A-88	Sequence 88, Appl1
19	231	36.3	557 4 US-09-338-933-88	Sequence 88, Appl1
20	231	36.3	557 4 US-09-215-681-88	Sequence 88, Appl1
21	231	36.3	581 4 US-09-601-144-67	Sequence 67, Appl1
22	231	36.3	594 4 US-09-404-879A-87	Sequence 87, Appl1
23	231	36.3	594 4 US-09-338-933-87	Sequence 87, Appl1
24	231	36.3	594 4 US-09-215-681-87	Sequence 87, Appl1
25	231	36.3	601 4 US-09-404-879A-133	Sequence 133, App
26	231	36.3	601 4 US-09-338-933-133	Sequence 133, App
27	231	36.3	601 4 US-09-215-681-133	Sequence 133, App
28	231	36.3	624 3 US-09-385-982-526	Sequence 526, App
29	231	36.3	630 3 US-08-180-371-5	Sequence 5, Appl1
30	231	36.3	630 3 PCR-US92-05707-5	Sequence 5, Appl1
31	230	36.1	630 3 US-08-180-371-17	Sequence 17, Appl1
32	223.5	35.1	318 4 US-09-313-294A-5868	Sequence 5868, Ap
33	216	33.9	605 3 US-09-385-982-485	Sequence 485, App
34	210	33.0	631 3 US-09-385-982-174	Sequence 174, App
35	206	32.3	914 1 US-08-386-729A-9	Sequence 9, Appl1
36	189	29.7	278 4 US-09-221-017B-373	Sequence 5840, Ap
37	178.5	28.0	7096 4 US-09-221-017B-373	Sequence 373, App
38	172.5	27.1	564 2 US-08-775-978-2	Sequence 2, Appl1
39	172	27.0	729 2 US-08-826-910-2	Sequence 2, Appl1
40	171.5	26.9	336 4 US-09-134-001C-1022	Sequence 1022, Ap
41	170	26.7	4403765 3 US-09-103-840A-2	Sequence 2, Appl1
42	170	26.7	4411529 3 US-09-103-840A-1	Sequence 1, Appl1
43	167.5	26.3	1230025 4 US-09-198-452A-1	Sequence 1, Appl1
44	161	25.3	438 4 US-09-328-352-1234	Sequence 1234, Ap
45	152.5	23.9	4969 1 US-08-260-582-76	Sequence 76, Appl1

## ALIGNMENTS

RESULT 1  
US-08-181-271A-103  
Sequence 103, Application US/08181271A  
Patent No. 5614395  
GENERAL INFORMATION:  
APPLICANT: Ryals, John A.  
APPLICANT: Alexander, Danny C.  
APPLICANT: Beck, James J.  
APPLICANT: Duesting, John H.  
APPLICANT: Friedrich, Leslie B.  
APPLICANT: Goodman, Robert M.  
APPLICANT: Harms, Christian  
APPLICANT: Meins, Jr., Frederick  
APPLICANT: Montoya, Alice  
APPLICANT: Moyer, Mary B.  
APPLICANT: Neuhaus, Jean-Marc  
APPLICANT: Payne, George B.  
APPLICANT: Sperison, Christoph  
APPLICANT: Stinson, Jeffrey R.  
APPLICANT: Uknes, Scott J.  
APPLICANT: Ward, Eric R.  
TITLE OF INVENTION: Williams, Sherioka C.  
TITLE OF INVENTION: CHEMICALLY REGULATABLE AND ANTI-PATHOGENIC  
NUMBER OF SEQUENCES: 106  
CORRESPONDENCE ADDRESSES:  
ADDRESSEE: CIBA-GEIGY Corporation  
STREET: 7 Skyline Drive  
CITY: Hawthorne  
STATE: New York  
COUNTRY: USA  
ZIP: 10532  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS



;; FILING DATE: 16-JUL-1993  
;; PRIOR APPLICATION DATA:  
;; APPLICATION NUMBER: US 07/937,197  
;; FILING DATE: 6-NOV-1992  
;; PRIOR APPLICATION DATA:  
;; APPLICATION NUMBER: US 07/678,378  
;; FILING DATE: 1-APR-1991  
;; PRIOR APPLICATION DATA:  
;; APPLICATION NUMBER: US 07/305,566  
;; FILING DATE: 6-FEB-1989  
;; PRIOR APPLICATION DATA:  
;; APPLICATION NUMBER: US 07/165,667  
;; FILING DATE: 8-MAR-1988  
;; PRIOR APPLICATION DATA:  
;; APPLICATION NUMBER: US 08/042,847  
;; FILING DATE: 6-APR-1993  
;; PRIOR APPLICATION DATA:  
;; APPLICATION NUMBER: US 07/632,441  
;; FILING DATE: 21-DEC-1990  
;; PRIOR APPLICATION DATA:  
;; APPLICATION NUMBER: US 07/425,504  
;; FILING DATE: 20-OCT-1989  
;; PRIOR APPLICATION DATA:  
;; APPLICATION NUMBER: US 07/848,506  
;; FILING DATE: 6-MAR-1992  
;; PRIOR APPLICATION DATA:  
;; APPLICATION NUMBER: US 07/768,122  
;; FILING DATE: 27-SEP-1991  
;; PRIOR APPLICATION DATA:  
;; APPLICATION NUMBER: US 07/580,431  
;; FILING DATE: 7-SEP-1990  
;; PRIOR APPLICATION DATA:  
;; APPLICATION NUMBER: US 07/368,672  
;; FILING DATE: 20-JUN-1989  
;; PRIOR APPLICATION DATA:  
;; APPLICATION NUMBER: US 07/329,018  
;; FILING DATE: 24-MAR-1989  
;; PRIOR APPLICATION DATA:  
;; APPLICATION NUMBER: US 08/045,957  
;; FILING DATE: 12-APR-1993  
;; ATTORNEY/AGENT INFORMATION:  
;; NAME: Elmer, James Scott  
;; REGISTRATION NUMBER: 36,129  
;; REFERENCE/DOCKET NUMBER: S-19825/P1/CGC 1727  
;; TELECOMMUNICATION INFORMATION:  
;; TELEPHONE: (919)541-8614  
;; TELEFAX: (919)541-8689  
;; INFORMATION FOR SEQ ID NO: 103:  
;; SEQUENCE CHARACTERISTICS:  
;; LENGTH: 653 base pairs  
;; TYPE: nucleic acid  
;; STRANDEDNESS: single  
;; TOPOLOGY: linear  
;; MOLECULE TYPE: DNA (genomic)  
;; US-08-449-315-103  
  
Alignment Scores:  
Pred. No.: 1,49e-35 Length: 653  
Score: 335.50 Matches: 65  
Percent Similarity: 72.13% Conservative: 23  
Best Local Similarity: 53.28% Mismatches: 30  
Query Match: 52.67% Indels: 4  
DB: 1 Gaps: 1  
  
US-09-786-715-2 (1-122) x US-08-449-315-103 (1-653)  
QY 2 AaagluglglglvalilleaacyshstlyslleaapglutrpgluglgluInleugly 21  
DB 72 TCCGAGGAGGAGGAGGTGTTCGCGTCCCAAGAGTTGAGAGATGGAAGAGTACTTCAAG 131  
QY 22 LysTrpLysAspSerGlnLysLeuValValAlaAspPheRhrAlaSerTrpCysGlyPro 41  
DB 132 AAAGCGTTGAGACTAAGAACTGCTGCTGATTTTACTGCTTCACTGATGTCGCGSCT 191

QY 42 CysArgAlaIleAlaIleProTyrPheThrRgIleuAlaIleLysAsnAsnProAsnValAlaIlePhe 61  
DB 192 TGCCGTTTATGTCGCCCAATTCCTGCTGACATTGCTAAAGATGCGCCCATGTTATATTC 251  
QY 62 LeuLysValAspValAspLysLeuAsnSerValAlaSer-LysTrpGluIleAsnAlaIle 81  
DB 252 CTCAGAGTTGATGTTGATGATGAAAGAGTGTTCAGCGGGAATGAGTGTGAGGCAT 311  
QY 81 tProThrPheValPheLeuLysLysGlyLysIleIleGluLysIleValAlaIleAspLys 101  
DB 312 GCCAACTTTGCTCTCATTAAGATGGAAGAAAGAGGAGAGAGTGTGTCGCAAGAA 371  
QY 101 sValGlyLeuSerLysLysIleLeuGluLeuSerGlyThrTrpProAlaIleThrSerth 121  
DB 372 AGAGGAGTTGACAGACCAACCATGATGTAAGCATGCTGCT-----CCTGTAAGTGTAC 422  
QY 121 rAla 122  
DB 423 TGCT 426  
  
RESULT 3  
US-08-444-803-103  
; Sequence 103, Application US/08444803  
; Patent No. 5654414  
; GENERAL INFORMATION:  
; APPLICANT: Ryals, John A.  
; APPLICANT: Alexander, Danny C.  
; APPLICANT: Beck, James J.  
; APPLICANT: Duesling, John H.  
; APPLICANT: Friedrich, Leslie B.  
; APPLICANT: Goodman, Robert M.  
; APPLICANT: Harms, Christian  
; APPLICANT: Meins, Jr., Frederick  
; APPLICANT: Montoya, Alice  
; APPLICANT: Moyer, Mary B.  
; APPLICANT: Neuhaus, Jean-Marc  
; APPLICANT: Payne, George B.  
; APPLICANT: Sperison, Christoph  
; APPLICANT: Stinson, Jeffrey R.  
; APPLICANT: Uknes, Scott J.  
; APPLICANT: Ward, Eric R.  
; APPLICANT: Williams, Shericca C.  
; TITLE OF INVENTION: CHEMICALLY REGULATABLE AND ANTI-PATHOGENIC  
; TITLE OF INVENTION: DNA SEQUENCES AND USES THEREOF  
; NUMBER OF SEQUENCES: 106  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: CIBA-GEIGY Corporation  
; STREET: 7 Skyline Drive  
; CITY: Hawthorne  
; STATE: New York  
; COUNTRY: USA  
; ZIP: 10532  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/444,803  
; FILING DATE: 19-MAY-1995  
; CLASSIFICATION: 536  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 08/181,271  
; FILING DATE: 13-JAN-94  
; APPLICATION NUMBER: US 08/093,301  
; FILING DATE: 16-JUL-1993  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 07/937,197  
; FILING DATE: 6-NOV-1992  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 07/678,378  
; FILING DATE: 1-APR-1991

PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/305,566  
FILING DATE: 6-FEB-1989  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/165,667  
FILING DATE: 8-MAR-1988  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/042,847  
FILING DATE: 6-APR-1993  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/632,441  
FILING DATE: 21-DEC-1990  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/425,504  
FILING DATE: 20-OCT-1989  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/848,506  
FILING DATE: 6-MAR-1992  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/768,122  
FILING DATE: 27-SEP-1991  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/580,431  
FILING DATE: 7-SEP-1990  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/368,672  
FILING DATE: 20-JUN-1989  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/329,018  
FILING DATE: 24-MAR-1989  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/045,957  
FILING DATE: 12-APR-1993  
ATTORNEY/AGENT INFORMATION:  
NAME: Elmer, James Scott  
REGISTRATION NUMBER: 36,129  
REFERENCE/DOCKET NUMBER: S-19825/P1/CGC 1727  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (919)541-8614  
TELEFAX: (919)541-8689  
INFORMATION FOR SEQ ID NO: 103:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 653 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: DNA (genomic)  
US-08-444-803-103

Alignment Scores:  
Pred. No.: 1,49e-35 Length: 653  
Score: 335.50 Matches: 65  
Percent Similarity: 72.13% Conservative: 23  
Best Local Similarity: 53.28% Mismatches: 30  
Query Match: 52.67% Indels: 4  
DB: 1 Gaps: 1

US-09-786-715-2 (1-122) x US-08-444-803-103 (1-653)

QY 2 AAlaGlUGlUGlInValIleAlaCySHsLysIleAspGluTrpGluGlnLeuGly 21  
Db 72 TCCGAGAGGAGCAAAATGTTGGCTGCCAACAGTTGAGCAAGAAAGCAAGTACTTCAG 131  
QY 22 LysTrpLysAspSerGluLysLeuValValValAspPheThrAlaSerTrpCysGlyPro 41  
Db 132 AAAGCGGTGAGGCTAGAAACAGTGAGTGGTGGATTTTACTCTTCATGAGTGGGSCCT 191  
QY 42 CysArgAlaIleAlaProTyrPheThrGluLeuAlaLysAsnAspProAsnValAlaPhe 61  
Db 192 TGGCGTTTATTTGCCCAATCTTCTGACATTCGTAAGAAAGATGCCCATGTATATTC 251  
QY 62 LeuLysValAspValAspGluLeuAsnSerValAlaSer-LysTrpGluIleAsnAlaLe 81

Db 252 CTCAGGTTGATGTGATGACGTGACACGCTTTACCGGGCAATGAGTGTGAGGCAAT 311  
QY 81 tProthrPheValPheLeuLysGlyLysIleIleGluLysIleValGlyAlaAspLys 101  
Db 312 GCCAAGCTTTGTTCTTCAATTAAAGTGGAAAAAGAGTGGACAGACTTGTGTCACAA 371  
QY 101 sValGlyLeuSerLysLysIleLeuGluLeuSerGlyTThrThrProAlaIleThrSer 121  
Db 372 AGAGAGTTCAGACAGACATAGTACGATGCTGCT-----CCTGCTACTGTCAC 422  
QY 121 rAla 122  
Db 423 TGCT 426

RESULT 4  
US-08-449-043-103  
Sequence 103, Application US/08449043  
Patent No. 5689044  
GENERAL INFORMATION:  
APPLICANT: Ryals, John A.  
APPLICANT: Alexander, Danny C.  
APPLICANT: Beck, James J.  
APPLICANT: Duesing, John H.  
APPLICANT: Friedrich, Leslie B.  
APPLICANT: Goodman, Robert M.  
APPLICANT: Harms, Christian  
APPLICANT: Melns, Jr., Frederick  
APPLICANT: Montoya, Alice  
APPLICANT: Moyer, Mary B.  
APPLICANT: Neuhaus, Jean-Marc  
APPLICANT: Payne, George B.  
APPLICANT: Speiserson, Christoph  
APPLICANT: Stinson, Jeffrey R.  
APPLICANT: Uknes, Scott J.  
APPLICANT: Ward, Eric R.  
APPLICANT: Williams, Shericca C.  
TITLE OF INVENTION: CHEMICALLY REGULATABLE AND ANTI-PATHOGENIC  
TITLE OF INVENTION: DNA SEQUENCES AND USES THEREOF  
NUMBER OF SEQUENCES: 106  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: CIBA-GEIGY Corporation  
STREET: 7 Skyline Drive  
CITY: Hawthorne  
STATE: New York  
COUNTRY: USA  
ZIP: 10532  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/449,043  
FILING DATE: 24-MAY-1995  
CLASSIFICATION: 800  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/181,271  
FILING DATE: 13-JAN-94  
APPLICATION NUMBER: US 08/093,301  
FILING DATE: 16-JUL-1993  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/937,197  
FILING DATE: 6-NOV-1992  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/678,378  
FILING DATE: 1-APR-1991  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/305,566  
FILING DATE: 6-FEB-1989  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/165,667  
FILING DATE: 8-MAR-1988  
PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 08/042,847  
FILING DATE: 6-APR-1993  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/632,441  
FILING DATE: 21-DEC-1990  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/425,504  
FILING DATE: 20-OCT-1989  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/848,506  
FILING DATE: 6-MAR-1992  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/768,122  
FILING DATE: 27-SEP-1991  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/580,431  
FILING DATE: 7-SEP-1990  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/368,672  
FILING DATE: 20-JUN-1989  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/329,018  
FILING DATE: 24-MAR-1989  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/045,957  
FILING DATE: 12-APR-1993  
ATTORNEY/AGENT INFORMATION:  
NAME: Elmer, James Scott  
REGISTRATION NUMBER: 36,129  
REFERENCE/DOCKET NUMBER: S-19825/P1/CGC 1727  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (919)541-8614  
TELEFAX: (919)541-8689  
INFORMATION FOR SEQ ID NO: 103:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 653 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: DNA (genomic)  
US-08-449-043-103

Alignment Scores:  
Pred. No.: 1,49e-35 Length: 653  
Score: 335.50 Matches: 65  
Percent Similarity: 72.13% Conservative: 23  
Best Local Similarity: 53.28% Mismatches: 30  
Query Match: 52.67% Indels: 4  
DB: 1 Gaps: 1

US-09-786-715-2 (1-122) x US-08-449-043-103 (1-653)

QY 2 Aatagluglglglnvallealacyshtslyslaspsglutrpeluglglngly 21  
:::|||||  
DB 72 TCCGAGGAGGCAAGTGTTCGCGCCACAGGTTGAGAGTGAACAGACTTCAG 131  
:::|||||  
QY 22 LystpLysAspSerGluLysLeuValValAlaspheThAlasertPrpGlyPro 41  
|||:::|||||  
DB 132 AAAGCGGTGAGACAACTGAGTGTGCGATTTTACGCTTCACAGTGCGGSCCT 191  
|||:::|||||  
QY 42 CysArGAlaLeAlProTyrPheThrgLuleAlaLysAsnAspProksnValAlape 61  
|||||:::|||||  
DB 192 TCCCGTTTATTCGCCCAATTTCTGCTGACATTGCTAAGAAAGATGCCCATGTTATTC 251  
|||||:::|||||  
QY 62 LeuLysValAspValAlaspLuleuAsnSerValAlaSer-LystpGluLleAsnAla 81  
|||||:::|||||  
DB 252 CTCAGGTTGATGATGAAGAACTGTTTCAGCGGAGAGTGAAGTGAGGCAAT 311  
|||||:::|||||  
QY 81 tProthrPheValPheLeuLysGlyLysLleleGluLysLleValGlyAlaAspLy 101  
|||||:::|||||  
DB 312 GCCAACCTTTGCTTCATTAAAGAGGAAAGAAAGTGTGTCAGCAAGAA 371  
|||||:::|||||  
QY 101 sValGlyLeuSerLysLysLleleuGluLeuSerGlyThrThrProAlaLethSerTh 121  
|||||:::|||||

DB 372 AGAGGAGTTGCAGACAGACCATCTAGTAAGATGCTGCT 11111111  
QY 121 rAla 122  
|||||  
DB 423 TCCT 426

RESULT 5  
US-08-456-265A-103  
Sequence 103, Application US/08456265A  
Patent No. 5767369  
GENERAL INFORMATION:  
APPLICANT: Alexander, Danny C.  
APPLICANT: Ryals, John A.  
APPLICANT: Goodman, Robert M.  
APPLICANT: Stinson, Jeffrey R.  
TITLE OF INVENTION: CHEMICALLY REGULATABLE AND ANTI-PATHOGENIC  
NUMBER OF SEQUENCES: 111  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: CIBA-GEIGY Corporation  
STREET: 520 White Plains Road, P.O. Box 2005  
CITY: Tarrytown  
STATE: New York  
COUNTRY: USA  
ZIP: 10591

COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/456,265A  
FILING DATE: 31-MAY-95  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/181,271  
FILING DATE: 13-JAN-1994  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/093,301  
FILING DATE: 16-JUL-1993  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/937,197  
FILING DATE: 6-NOV-1992  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/678,378  
FILING DATE: 1-APR-1991  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/305,566  
FILING DATE: 6-FEB-1989  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/165,667  
FILING DATE: 8-MAR-1988  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/042,847  
FILING DATE: 6-APR-1993  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/632,441  
FILING DATE: 21-DEC-1990  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/425,504  
FILING DATE: 20-OCT-1989  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/848,506  
FILING DATE: 6-MAR-1992  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/768,122  
FILING DATE: 27-SEP-1991  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/580,431  
FILING DATE: 7-SEP-1990  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/368,672

FILING DATE: 20-JUN-1989  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/329,018  
FILING DATE: 24-MAR-1989  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/045,957  
FILING DATE: 12-APR-1993  
ATTORNEY/AGENT INFORMATION:  
NAME: Meigs, J. Timothy  
REGISTRATION NUMBER: 38,241  
REFERENCE/DOCKET NUMBER: S-19825/PL/CGC 1727/DIV10  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (919)541-8367  
TELEFAX: (919)541-8689  
INFORMATION FOR SEQ ID NO: 103:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 653 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: DNA (genomic)  
US-08-456-265A-103

Alignment Scores:  
Pred. No.: 1,49e-35 Length: 653  
Score: 335.50 Matches: 65  
Percent Similarity: 72.13% Conservative: 23  
Best Local Similarity: 53.28% Mismatches: 30  
Query Match: 52.67% Indels: 4  
DB: 1 Gaps: 1

US-09-786-715-2 (1-122) x US-08-456-265A-103 (1-653)

QY 2 AAlAGlUGlUGlGlnValIleAlaCySHlSLyslleAspGluTrpGluGlnLeuGly 21  
Db 72 TCCGAGGAGGAGCAAGGTTCCGCTGCCACAGGTTGAGGAATGGAACGACTTCAAG 131  
QY 22 LysTrpLysAspSerGluLysLeuValValAlaSpheThrAlaSerTrpCysGlyPro 41  
Db 132 AAAGCGCTGAGACTAGAAACGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 191  
QY 42 CysArgAlaIleAlaProTrpPheThrGlnLeuAlaLysAsnAsnProAsnValAlaPhe 61  
Db 192 TGCCGCTTATTCCTCCCAATTCCTGCTGACATTCCTAAGAAAGATGTTATTTTC 251  
QY 62 LeuLysValAspValAspGluLeuAsnSerValAlaSerLysTrpGluLeuAsnAla 81  
Db 252 CTCAGGTTGATGTTGATGAACTGAACTGTTTCAACGGGAATGGAGTGGAGCAAT 311  
QY 81 tProThrPheValPheLeuLysLysGlyLysIleIleGluLysIleValGlyAlaAspLys 101  
Db 312 GCCAAGCTTTGCTCTCATTAAGATGGAAGAAAGATGACAGATGCTGTCGCAAGAA 371  
QY 101 sValGlyLeuSerLysLysIleLeuGluLeuSerGlyThrThrProAlaIleAsnSer 121  
Db 372 AGAGGAGTTCACACAGACCATTAAGTGAAGATGCTGCT-----CCTGCTACTGTCAC 422  
QY 121 rAla 122  
Db 423 TGCT 426

RESULT 6  
US-08-455-416-103  
Sequence 103, Application US/08455416  
Patent No. 5777200  
GENERAL INFORMATION:  
APPLICANT: Ryals, John A.  
APPLICANT: Alexander, Danny C.  
APPLICANT: Beck, James J.  
APPLICANT: Duesing, John H.  
APPLICANT: Friedlich, Leslie B.  
APPLICANT: Goodman, Robert M.  
APPLICANT: Harms, Christian

APPLICANT: Meigs, Jr., Frederick  
APPLICANT: Montoya, Alice  
APPLICANT: Moyer, Mary B.  
APPLICANT: Neuhaus, Jean-Marc  
APPLICANT: Payne, George B.  
APPLICANT: Sperlson, Christoph  
APPLICANT: Stinson, Jeffrey R.  
APPLICANT: Uknes, Scott J.  
APPLICANT: Ward, Eric R.  
APPLICANT: Williams, Shericea C.  
TITLE OF INVENTION: CHEMICALLY REGULATABLE AND ANTI-PATHOGENIC  
NUMBER OF SEQUENCES: 106  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: CIBA-GEIGY Corporation  
STREET: 7 Skyline Drive  
CITY: Hawthorne  
STATE: New York  
COUNTRY: USA  
ZIP: 10532  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/455,416  
FILING DATE: 31-MAY-1995  
CLASSIFICATION: 800  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/181,271  
FILING DATE: 13-JAN-94  
APPLICATION NUMBER: US 08/093,301  
FILING DATE: 16-JUL-1993  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/937,197  
FILING DATE: 6-NOV-1992  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/678,378  
FILING DATE: 1-APR-1991  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/305,566  
FILING DATE: 6-FEB-1989  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/165,667  
FILING DATE: 8-MAR-1988  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/042,847  
FILING DATE: 6-APR-1993  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/632,441  
FILING DATE: 21-DEC-1990  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/425,504  
FILING DATE: 20-OCT-1989  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/848,506  
FILING DATE: 6-MAR-1992  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/768,122  
FILING DATE: 27-SEP-1991  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/580,431  
FILING DATE: 7-SEP-1990  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/368,672  
FILING DATE: 20-JUN-1989  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/329,018  
FILING DATE: 24-MAR-1989  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/045,957  
FILING DATE: 12-APR-1993

ATTORNEY/AGENT INFORMATION:  
NAME: Elmer, James Scott  
REGISTRATION NUMBER: 36,129  
REFERENCE/DOCKET NUMBER: S-19825/P1/CCG 1727  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (919)541-8614  
TELEFAX: (919)541-8689  
INFORMATION FOR SEQ ID NO: 103:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 653 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: DNA (genomic)  
US-08-455-416-103

Alignment Scores:  
Pred. No.: 1.49e-35 Length: 653  
Score: 335.50 Matches: 65  
Percent Similarity: 72.13% Conservative: 23  
Best Local Similarity: 53.28% Mismatches: 30  
Query Match: 52.67% Indels: 4  
DB: 1 Gaps: 1

US-09-786-715-2 (1-122) x US-08-455-416-103 (1-653)

QY 2 AAlagluGluglYglVallleAlacysHslsYlslleAspGluTrpGluInleugly 21  
Db 72 TCCGAGAGGAGGACCAAGTGTGCGCTGCCACAGAGTTGAGCAATGGACAGACTTCAAG 131  
QY 22 lYstrPlYaspSerserlYlslleuValValAlaspPheThrAlaserTrpCysglYPro 41  
Db 133 AAGGCGGTGAGACTAAGAACTGGTGGTGGATTTCATGCTTCATGCGGSCCT 191  
QY 42 CysatgAlaIlleAlaProTyrPheThrGluLeuAlaYlslslsAsnProAsnValAlaPhe 61  
Db 192 TCCCGTTTATGCCCCAATCTCTGCTGACATGCTAAGAGATGCCCATGTTATATTC 251  
QY 62 lLeuYValAspValAspGluLeuAsnserValAlaser-LYstrPgluIlleAsnAlaMe 81  
Db 252 CTCAGAGTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 311  
QY 81 lProThrPheValPheLeuYlslYlslleleGluYlslleValAlaAlaAspY 101  
Db 312 GCCAAGTTTGTCTTCATTAAGAGATGAGAAAGAGAGAGAGAGAGAGAGAGAGAGAGAG 371  
QY 101 sValglYleuSerYlslYlslleuGluLeuSerGlyThrProAlaAlaThrSerTh 121  
Db 372 AGAGGAGTTGCAGAGACCATAGTGAAGCATGCTGCT-----CCTGCTACTGTGAC 422  
QY 121 rAla 122  
Db 423 TCCT 426

RESULT 7  
US-08-455-244-103  
Sequence 103, Application US/08455244  
Patent No. 5789214  
GENERAL INFORMATION:  
APPLICANT: Ryals, John A.  
APPLICANT: Alexander, Danny C.  
APPLICANT: Beck, James J.  
APPLICANT: Duesing, John H.  
APPLICANT: Friedrich, Leslie B.  
APPLICANT: Goodman, Robert M.  
APPLICANT: Harms, Christian  
APPLICANT: Meins, Jr., Frederick  
APPLICANT: Montoya, Alice  
APPLICANT: Moyer, Mary B.  
APPLICANT: Neuhaus, Jean-Marc  
APPLICANT: Payne, George B.  
APPLICANT: Sperison, Christoph  
APPLICANT: Stinson, Jeffrey R.

APPLICANT: Uknes, Scott J.  
APPLICANT: Ward, Eric R.  
APPLICANT: Williams, Shericea C.  
TITLE OF INVENTION: CHEMICALLY REGULATABLE AND ANTI-PATHOGENIC  
TITLE OF INVENTION: DNA SEQUENCES AND USES THEREOF  
NUMBER OF SEQUENCES: 106  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: CIBA-GEIGY Corporation  
STREET: 7 Skyline Drive  
CITY: Hawthorne  
STATE: New York  
COUNTRY: USA  
ZIP: 10532

COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/455,244  
FILING DATE: 31-MAY-1995  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/181,271  
FILING DATE: 13-JAN-94  
APPLICATION NUMBER: US 08/093,301  
FILING DATE: 16-JUL-1993  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/937,197  
FILING DATE: 6-NOV-1992  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/678,378  
FILING DATE: 1-APR-1991  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/305,566  
FILING DATE: 6-FEB-1989  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/165,667  
FILING DATE: 8-MAR-1988  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/042,847  
FILING DATE: 6-APR-1993  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/632,441  
FILING DATE: 21-DEC-1990  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/425,504  
FILING DATE: 20-OCT-1989  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/848,506  
FILING DATE: 6-MAR-1992  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/768,122  
FILING DATE: 27-SEP-1991  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/580,431  
FILING DATE: 7-SEP-1990  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/368,672  
FILING DATE: 20-JUN-1989  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/329,018  
FILING DATE: 24-MAR-1989  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/045,957  
FILING DATE: 12-APR-1993  
ATTORNEY/AGENT INFORMATION:  
NAME: Elmer, James Scott  
REGISTRATION NUMBER: 36,129  
REFERENCE/DOCKET NUMBER: S-19825/P1/CCG 1727  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (919)541-8614  
TELEFAX: (919)541-8689

INFORMATION FOR SEQ ID NO: 103:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 653 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: DNA (genomic)  
US-08-455-244-103

Alignment Scores:  
Pred. No.: 1.49e-35 Length: 653  
Score: 335.50 Matches: 65  
Percent Similarity: 72.13% Conservative: 23  
Best Local Similarity: 53.28% Mismatches: 30  
Query Match: 52.67% Indels: 4  
DB: 1 Gaps: 1

US-09-786-715-2 (1-122) x US-08-455-244-103 (1-653)

QY 2 ALaGluGluGluGluValIleAlaCysHisLysIleAspGluTrpGluGluLeuGly 21  
:::|||||  
DB 72 TCCGAGAGAGGACAGGTTCCGCTGCCACAGGTTGAGATGAGACGACTTCAAG 131  
|||  
QY 22 LysTrpLysAspSerGluLysLeuValValAlaAspPheThrAlaSerTrpCysGlyPro 41  
|||  
DB 132 AAAGCGCTGAGACTAGAAAGAACTGGTGCTGCTGCTTACTGCTTCATGGTGGGSCCT 191  
|||  
QY 42 CysArgAlaIleAlaProTyrPheThrGluLeuAlaLysAsnAsnProAsnValAlaPhe 61  
|||||  
DB 192 TGGCGTTTATGCTCCCAATCTTGTGACATTCCTAAGATGAGATGATATATTC 251  
|||  
QY 62 LeuLysValAspValaAspGluLeuAsnSerValAlaSer-LysTrpGluIleAsnAlaMe 81  
|||||  
DB 252 CTCAGGTTGATGTTGATGAAGTCAAGACTGTTTACCGGGAATGAGAGTGGAGGCAAT 311  
|||  
QY 81 tProThPheValaPheLeuLysGlyLysIleIleGluLysIleValGlyAlaAspLy 101  
|||||  
DB 312 GCCAACTTTGCTCTCAATTAAGATGAAAAAGAGTGACAGACTGTTGTCGCAAGAA 371  
|||  
QY 101 sValGlyLeuSerLysLysIleLeuGluLeuSerGlyThrThrProAlaIleThrSer 121  
|||  
DB 372 AGAGAGATTGACAGCACCATAGTGAAGCATGCTGCT-----CCTGCTACTGTAC 422  
|||  
QY 121 Ala 122  
|||  
DB 423 TGCT 426

RESULT 8  
US-08-454-876-103  
Sequence 103, Application US/08454876  
Patent No. 5804693  
GENERAL INFORMATION:  
APPLICANT: Ryals, John A.  
APPLICANT: Alexander, Danny C.  
APPLICANT: Beck, James J.  
APPLICANT: Duesing, John H.  
APPLICANT: Friedlich, Leslie B.  
APPLICANT: Goodman, Robert M.  
APPLICANT: Harms, Christian  
APPLICANT: Meins, Jr., Frederick  
APPLICANT: Montoya, Alice  
APPLICANT: Moyer, Mary B.  
APPLICANT: Neuhaus, Jean-Marc  
APPLICANT: Payne, George B.  
APPLICANT: Sperison, Christoph  
APPLICANT: Stinson, Jeffrey R.  
APPLICANT: Uknes, Scott J.  
APPLICANT: Ward, Eric R.  
TITLE OF INVENTION: CHEMICALS C.  
TITLE OF INVENTION: CHEMICALLY REGULATABLE AND ANTI-PATHOGENIC  
NUMBER OF SEQUENCES: 106  
CORRESPONDENCE ADDRESS:

ADDRESSEE: CIBA-GEIGY Corporation  
STREET: 7 Skyline Drive  
CITY: Hawthorne  
STATE: New York  
COUNTRY: USA  
ZIP: 10532  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/454,876  
FILING DATE: 31-MAY-1995  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/181,271  
FILING DATE: 13-JAN-94  
APPLICATION NUMBER: US 08/093,301  
FILING DATE: 16-JUL-1993  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/937,197  
FILING DATE: 6-NOV-1992  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/678,378  
FILING DATE: 1-APR-1991  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/305,566  
FILING DATE: 6-FEB-1989  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/165,667  
FILING DATE: 8-MAR-1988  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/042,847  
FILING DATE: 6-APR-1993  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/632,441  
FILING DATE: 21-DEC-1990  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/425,504  
FILING DATE: 20-OCT-1989  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/848,506  
FILING DATE: 6-MAR-1992  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/768,122  
FILING DATE: 27-SEP-1991  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/580,431  
FILING DATE: 7-SEP-1990  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/368,672  
FILING DATE: 20-JUN-1989  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/329,018  
FILING DATE: 24-MAR-1989  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/045,957  
FILING DATE: 12-APR-1993  
ATTORNEY/AGENT INFORMATION:  
NAME: Elmer, James Scott  
REGISTRATION NUMBER: 36,129  
REFERENCE/DOCKET NUMBER: S-19825/P1/CGC 1727  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (919)541-8614  
TELEFAX: (919)541-8689  
INFORMATION FOR SEQ ID NO: 103:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 653 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: DNA (genomic)



US-08-454-876-103

## Alignment Scores:

Pred. No.:	1.49e-35	Length:	653
Score:	335.50	Matches:	65
Percent Similarity:	72.13%	Conservative:	23
Best Local Similarity:	53.28%	Mismatches:	30
Query Match:	52.67%	Indels:	4
DB:	1	Gaps:	1

US-09-786-715-2 (1-122) x US-08-454-876-103 (1-653)

```
QY      2 Alagluglgiyginvalilaleaacyshslslyslleapglutprgluglincly 21
DB      72 TCCGAGGAGGACAAGTGTGCGGCGCCACAGGTGAGATGGAAGAGTCTCAAG 131
QY      22 LysTrpLysaspseglulysleuValValValaspphetrAlaSerTrpCysGlyPro 41
DB      132 AAAGCGCTTGAGACTAAGAAACTGGTGGTGTGCTGATTTACTGCTCAATGGCGGSCCT 191
QY      42 CysArGalaIleAlaProTyPheThrGluleuAlaLysAsnAsnProAsnValAlaPhe 61
DB      192 TCCGCTTTATTTGCCCAATCTCTGCTACATGCTAAGAAAGTCCCACTTATATTC 251
QY      62 LeulysValaspValaspGluleuAsnSerValAlaSer-LysTrpGlulileaspAlaMe 81
DB      252 CTCAGAGTTGATGTTGATGAAGTGAAGACTGTTTCAGCGGGAAGTGAAGTGGAGGCAAT 311
QY      81 tProThrPheValPheLeuLysLysGlyLysIleleuLysIleValGlyAlaAspLys 101
DB      312 GCCAAGCTTCTCTTCATTAAGAAAGTGAAGTGAAGAGTGTGGTGGCCAGAA 371
QY      101 sValGlyLeuSerLysLysIleleuLysSerGlyThrProAlaAlaThSerTh 121
DB      372 AGAGAGTTGACGACGACCATGATGAGCATGCTGCT-----CCTGCTACTGTAC 422
QY      121 rAla 122
DB      423 TGCT 426
```

## RESULT 9

US-08-457-364-103  
Sequence 103, Application US/08457364  
Patent No. 5847258

## GENERAL INFORMATION:

APPLICANT: Ryals, John A.  
APPLICANT: Alexander, Danny C.  
APPLICANT: Beck, James J.  
APPLICANT: Duesing, John H.  
APPLICANT: Friedrich, Leslie B.  
APPLICANT: Goodman, Robert M.  
APPLICANT: Harms, Christian  
APPLICANT: Meins, Jr., Frederick  
APPLICANT: Montoya, Alice  
APPLICANT: Moyer, Mary B.  
APPLICANT: Neuhaus, Jean-Marc  
APPLICANT: Payne, George B.  
APPLICANT: Sperison, Christoph  
APPLICANT: Stinson, Jeffrey R.  
APPLICANT: Uknes, Scott J.  
APPLICANT: Ward, Eric R.  
APPLICANT: Williams, Shericca C.  
TITLE OF INVENTION: CHEMICALLY REGULATABLE AND ANTI-PATHOGENIC  
TITLE OF INVENTION: DNA SEQUENCES AND USES THEREOF  
NUMBER OF SEQUENCES: 106  
CORRESPONDENCE ADDRESSES:  
ADDRESS: CIBA-GEIGY Corporation  
STREET: 7 Skyline Drive  
CITY: Hawthorne  
STATE: New York  
COUNTRY: USA  
ZIP: 10532

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/457,364  
FILING DATE: 31-MAY-1995  
CLASSIFICATION: 800  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/181,271  
FILING DATE: 13-JAN-94  
APPLICATION NUMBER: US 08/093,301  
FILING DATE: 16-JUL-1993  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/937,197  
FILING DATE: 6-NOV-1992  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/678,378  
FILING DATE: 1-APR-1991  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/305,566  
FILING DATE: 6-FEB-1989  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/165,667  
FILING DATE: 8-MAR-1988  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/042,847  
FILING DATE: 6-APR-1993  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/632,441  
FILING DATE: 21-DEC-1990  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/425,504  
FILING DATE: 20-OCT-1989  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/848,506  
FILING DATE: 6-MAR-1992  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/768,122  
FILING DATE: 27-SEP-1991  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/580,431  
FILING DATE: 7-SEP-1990  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/368,672  
FILING DATE: 20-JUN-1989  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/329,018  
FILING DATE: 24-MAR-1989  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/045,957  
FILING DATE: 12-APR-1993  
ATTORNEY/AGENT INFORMATION:  
NAME: Elmer, James Scott  
REGISTRATION NUMBER: 36,129  
REFERENCE/DOCKET NUMBER: S-19825/P1/CGC 1727  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (919)541-8614  
TELEFAX: (919)541-8689  
INFORMATION FOR SEQ ID NO: 103:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 653 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: DNA (genomic)  
US-08-457-364-103

Alignment Scores:

Pred. No.:	1.49e-35	Length:	653
Score:	335.50	Matches:	65
Percent Similarity:	72.13%	Conservative:	23
Best Local Similarity:	53.28%	Mismatches:	30



Db 72 TCCGAGGAGGCAAGTGTTCGGCTGCCCAAGGTTGAGGAATGGAACGAGTACTTCAG 131  
QY 22 LysTrpLysAspSerGluLysLeuValValAlaSpheThAlaSerTrpCysGlyPro 41  
Db 132 AAAGCGCTTGAGACTAGAACTGGTGGTGCGATTTCCTTCATGCTGCGGSCCT 191  
QY 42 CysArgAlaIleAlaProTyrPheThrGluLeuAlaLysAsnAsnProAsnValAlaPhe 61  
Db 192 TCCCGCTTTATGCCCCCATTTCTGTGACATTTGCTAAGAAAGATGCCCATGTTATATTC 251  
QY 62 LeuLysValAspValAlaSpGluLeuAsnSerValAlaSer-LysTrpGluIleAsnAlaPhe 81  
Db 252 CTCAGGCTTGATGTGATGCACTGTAAGACTGTTTCACGCGGAATGAGTGTGAGGCAAT 311  
QY 81 ProThrPheValPheLeuLysLysGlyLysIleIleGluLysIleValAlaAspLys 101  
Db 312 GCCAACTTTTGTCTTCATTTAAAGATGGAAGAAAGAGGACAGATGTTGGGCCAAGA 371  
QY 101 sValGlyLeuSerLysLysLysLysLeuGluLeuSerGlyThrThrProAlaAlaThrSerTh 121  
Db 372 AGAGAGTTGCGACGACCATGATGTAGACATGCTGCT-----CCTGCTACTGTCAC 422  
QY 121 PAla 122  
Db 423 TGCT 426

## RESULT 11

US-08-456-240-103  
Sequence 103, Application US/08456240

Patent No. 5856154

GENERAL INFORMATION:

APPLICANT: Ryals, John A.

APPLICANT: Alexander, Danny C.

APPLICANT: Beck, James J.

APPLICANT: Duesting, John H.

APPLICANT: Friedrich, Leslie B.

APPLICANT: Goodman, Robert M.

APPLICANT: Harms, Christian

APPLICANT: Meins, Jr., Frederick

APPLICANT: Montoya, Alice

APPLICANT: Moyer, Mary B.

APPLICANT: Neuhaus, Jean-Marc

APPLICANT: Payne, George B.

APPLICANT: Sperison, Christoph

APPLICANT: Stinson, Jeffrey R.

APPLICANT: Uknes, Scott J.

APPLICANT: Ward, Eric R.

APPLICANT: Williams, SheriLeca C.

TITLE OF INVENTION: CHEMICALLY REGULATABLE AND ANTI-PATHOGENIC

TITLE OF INVENTION: DNA SEQUENCES AND USES THEREOF

NUMBER OF SEQUENCES: 106

CORRESPONDENCE ADDRESS:

ADDRESSEE: CIBA-GEIGY Corporation

STREET: 7 Skyline Drive

CITY: Hawthorne

STATE: New York

COUNTRY: USA

ZIP: 10532

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patent Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/456,240

FILING DATE: 31-MAY-1995

CLASSIFICATION: 800

PRIOR APPLICATION DATA:

APPLICATION NUMBER: 08/181,271

FILING DATE: 13-JAN-94

APPLICATION NUMBER: US 08/093,301

FILING DATE: 16-JUL-1993

PRIOR APPLICATION DATA:

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 07/937,197  
FILING DATE: 6-NOV-1992  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/678,378  
FILING DATE: 1-APR-1991  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/305,566  
FILING DATE: 6-FEB-1989  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/165,667  
FILING DATE: 8-MAR-1988  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/042,847  
FILING DATE: 6-APR-1993  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/632,441  
FILING DATE: 21-DEC-1990  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/425,504  
FILING DATE: 20-OCT 1989  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/848,506  
FILING DATE: 6-MAR-1992  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/768,122  
FILING DATE: 27-SEP-1991  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/580,431  
FILING DATE: 7-SEP-1990  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/368,672  
FILING DATE: 20-JUN-1989  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/329,018  
FILING DATE: 24-MAR-1989  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/045,957  
FILING DATE: 12-APR-1993  
ATTORENEY/AGENT INFORMATION:  
NAME: Elmer, James Scott  
REGISTRATION NUMBER: 36,129  
REFERENCE/DOCKET NUMBER: S-19825/P1/CGC 1727  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (919)541-8614  
TELEFAX: (919)541-8689  
INFORMATION FOR SEQ ID NO: 103:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 653 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: DNA (genomic)  
US-08-456-240-103

Alignment Scores:  
Pred. No.: 1,49e-35 Length: 653  
Score: 335.50 Matches: 65  
Percent Similarity: 72.138 Conservative: 23  
Best Local Similarity: 53.288 Mismatches: 30  
Query Match: 52.67% Indels: 4  
Gaps: 1

US-09-786-715-2 (1-122) x US-08-456-240-103 (1-653)

QY 2 AlaGluGluGlyGlnValIleAlaCysHisLysIleAspGluTrpGluGlnLeuGly 21  
Db 72 TCCGAGGAGGCAAGTGTTCGGCTGCCCAAGGTTGAGGAATGGAACGAGTACTTCAG 131  
QY 22 LysTrpLysAspSerGluLysLeuValValAlaSpheThAlaSerTrpCysGlyPro 41  
Db 132 AAAGCGCTTGAGACTAGAACTGGTGGTGCGATTTCCTTCATGCTGCGGSCCT 191  
QY 42 CysArgAlaIleAlaProTyrPheThrGluLeuAlaLysAsnAsnProAsnValAlaPhe 61

```
Db      192 TCCCGTTTATATCCCAATTCCTTGTGACATTGCTAAGAAGATGCCCATGTTATATTC 251
Qy      62 LeuLyValAspValAspGluLeuAsnSerValAlaSer-LysTrpGluIleAsnAlaMe 81
Db      252 CTCAGGTTGATGTGATGAGTGAAGTGAAGTGTTCACGGGAATGGAGTGGAGGCAAT 311
Qy      81 tProthrPheValPheLeuLysLysGlyLysIleIleGluLysIleValGlyAlaAspLy 101
Db      312 GCCAAGTTTGTCTTCAATTAAGATGGAAGAAAGAGTGTGTCGCCAAGAA 371
Qy      101 sValGlyLeuSerLysLysIleLeuGluLeuSerGlyThrThrProAlaAlaThrSerTh 121
Db      372 AGAGGAGTTGCACAGACCATAGTGAAGCATGCTGCT-----CCTGCTACTGTAC 422
Qy      121 rAla 122
Db      423 TGCT 426

RESULT 12
US-08-455-736-103
; Sequence 103, Application US/08455736
; Patent No. 5880328
; GENERAL INFORMATION:
; APPLICANT: Ryals, John A.
; APPLICANT: Alexander, Danny C.
; APPLICANT: Beck, James J.
; APPLICANT: Duesing, John H.
; APPLICANT: Friedrich, Leslie B.
; APPLICANT: Goodman, Robert M.
; APPLICANT: Harms, Christian
; APPLICANT: Meins, Jr., Frederick
; APPLICANT: Montoya, Alice
; APPLICANT: Moyer, Mary B.
; APPLICANT: Neuhaus, Jean-Marc
; APPLICANT: Payne, George B.
; APPLICANT: Sperison, Christoph
; APPLICANT: Stinson, Jeffrey R.
; APPLICANT: Uknes, Scott J.
; APPLICANT: Ward, Eric R.
; APPLICANT: Williams, Shericea C.
; TITLE OF INVENTION: CHEMICALLY REGULATABLE AND ANTI-PATHOGENIC
; TITLE OF INVENTION: DNA SEQUENCES AND USES THEREOF
; NUMBER OF SEQUENCES: 106
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: CIBA-GEIGY Corporation
; STREET: 7 Skyline Drive
; CITY: Hawthorne
; STATE: New York
; COUNTRY: USA
; ZIP: 10532
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/455,736
; FILING DATE: 31-MAY-1995
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/181,271
; FILING DATE: 13-JAN-1994
; APPLICATION NUMBER: US 08/093,301
; FILING DATE: 16-JUL-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/937,197
; FILING DATE: 6-NOV-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/678,378
; FILING DATE: 1-APR-1991
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/305,566

; FILING DATE: 6-FEB-1989
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/165,667
; FILING DATE: 8-MAR-1988
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/042,847
; FILING DATE: 6-APR-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/632,441
; FILING DATE: 21-DEC-1990
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/425,504
; FILING DATE: 20-OCT-1989
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/848,506
; FILING DATE: 6-MAR-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/768,122
; FILING DATE: 27-SEP-1991
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/580,431
; FILING DATE: 7-SEP-1990
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/368,672
; FILING DATE: 20-JUN-1989
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/329,018
; FILING DATE: 24-MAR-1989
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/045,957
; FILING DATE: 12-APR-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: Elmer, James Scott
; REGISTRATION/DOCKET NUMBER: 36,129
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (919)541-8614
; TELEFAX: (919)541-8689
; INFORMATION FOR SEQ ID NO: 103:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 653 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; US-08-455-736-103

Alignment Scores:
Pred. No.: 1,49e-35 Length: 653
Score: 335.50 Matches: 65
Percent Similarity: 72.13% Conservative: 23
Best Local Similarity: 53.28% Mismatches: 30
Query Match: 52.67% Indels: 4
                Gaps: 1

US-09-786-715-2 (1-122) x US-08-455-736-103 (1-653)
Qy      2 AlaGluGluGlyValIleAlaGlySHsLysIleAspGluTrpGluGlyLeuGly 21
Db      72 TCCGAGAGGAGGACAAGTGTTCGCGCTGCCACAAAGTTGAGGAATGGAACGACTCTCAAG 131
Qy      22 LysTrpLysAspSerGluLysLeuValIleValAspPheThrAlaSerTrpCysGlyPro 41
Db      132 AAGGCGTTGAGACTAAGAAGACGTGGTGTGATTTACTGCTTCATGTCGGGSCCT 191
Qy      42 CysArgAlaIleAlaProGlyThrGluLeuAlaLysAsnAsnProAsnValAlaPhe 61
Db      192 TCCCGTTTATATCCCAATTCCTTGTGACATTGCTAAGAAGATGCCCATGTTATATTC 251
Qy      62 LeuLyValAspValAspGluLeuAsnSerValAlaSer-LysTrpGluIleAsnAlaMe 81
Db      252 CTCAGGTTGATGTGATGAGTGAAGTGAAGTGTTCACGGGAATGGAGTGGAGGCAAT 311
```











GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - nucleic search, using frame\_plus\_p2n model

Run on: August 17, 2003, 19:25:47 ; Search time 142.207 Seconds  
(without alignments)  
1919.938 Million cell updates/sec

Title: US-09-786-715-2

Perfect score: 637  
Sequence: 1 MAERGQYIACHKIDWEGOL.....GLSKILELSTGPATSTA 122

Scoring table:  
BLOSUM62  
Xgapop 10.0, Xgapext 0.5  
Ygapop 10.0, Ygapext 0.5  
Fgapop 6.0, Fgapext 7.0  
Delop 6.0, Delext 7.0

Searched: 1504479 seqs, 1118970152 residues

Total number of hits satisfying chosen parameters: 3008958

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Command line parameters:

-MODEL=frame+g2n.model -DEV=xlp  
-O=/cgn2\_1/USPTO.spool/US09786715/unal\_11082003\_150515\_6129/app\_query.fasta\_1\_1052  
-DB=Published.Applications\_NA -OFMT=fastap -SUFFIX=rnpb -MINMATCH=0.1  
-LOOCL=0 -LOOEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blomsu62  
-TRANS=human40.cdi -LIST=45 -DOCALIGN=200 -THR SCORE=pcot -THR MAX=100  
-THR MIN=0 -ALIGN=15 -MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINTEN=0  
-MAXLEN=2000000000 -USER=US09786715.eccn\_1\_1\_560.etrunal\_11082003\_150515\_6129  
-NCPU=6 -ICPU=3 -NO\_MAP -LARGEQUERY -NEG\_SCORES=0 -WAIT -DSPLOCK=100  
-LONLOG -DEV.TIMEOUT=120 -WARN.TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5  
-FGAPOP=6 -FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : Published Applications\_NA:\*

1: /cgn2\_6/ptodata/2/pubpna/US07\_PUBCOMB.seq:\*  
2: /cgn2\_6/ptodata/2/pubpna/PTC\_NEW\_PUB.seq:\*  
3: /cgn2\_6/ptodata/2/pubpna/US06\_NEW\_PUB.seq:\*  
4: /cgn2\_6/ptodata/2/pubpna/US06\_PUBCOMB.seq:\*  
5: /cgn2\_6/ptodata/2/pubpna/US07\_NEW\_PUB.seq:\*  
6: /cgn2\_6/ptodata/2/pubpna/PCITUS\_PUBCOMB.seq:\*  
7: /cgn2\_6/ptodata/2/pubpna/US08\_NEW\_PUB.seq:\*  
8: /cgn2\_6/ptodata/2/pubpna/US08\_PUBCOMB.seq:\*  
9: /cgn2\_6/ptodata/2/pubpna/US09A\_PUBCOMB.seq:\*  
10: /cgn2\_6/ptodata/2/pubpna/US09C\_PUBCOMB.seq:\*  
11: /cgn2\_6/ptodata/2/pubpna/US09C\_NEW\_PUB.seq:\*  
12: /cgn2\_6/ptodata/2/pubpna/US09A\_PUBCOMB.seq:\*  
13: /cgn2\_6/ptodata/2/pubpna/US10A\_PUBCOMB.seq:\*  
14: /cgn2\_6/ptodata/2/pubpna/US10B\_PUBCOMB.seq:\*  
15: /cgn2\_6/ptodata/2/pubpna/US60\_NEW\_PUB.seq:\*  
16: /cgn2\_6/ptodata/2/pubpna/US60\_PUBCOMB.seq:\*  
17: /cgn2\_6/ptodata/2/pubpna/US60\_PUBCOMB.seq:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	377	59.2	3888	9	US-09-897-898-10 Sequence 10, Appl

2	376	59.0	345	10	US-09-938-842A-163	Sequence 163, App
3	376	59.0	3129	9	US-09-897-898-5	Sequence 5, Appl
4	376	59.0	3888	9	US-09-897-898-7	Sequence 7, Appl
5	360.5	56.6	402	10	US-09-878-574-3723	Sequence 3723, Ap
6	354	55.6	392	10	US-09-878-574-104	Sequence 104, App
7	345.5	54.2	360	10	US-09-938-842A-164	Sequence 164, App
8	343	53.8	390	10	US-09-878-574-1462	Sequence 1462, Ap
9	343	53.8	540	12	US-10-349-782-12	Sequence 12, Appl
10	323	50.7	357	10	US-09-938-842A-1736	Sequence 1736, Ap
11	323	50.7	357	14	US-10-323-362-1	Sequence 1, Appl
12	314	49.3	659	12	US-10-349-782-9	Sequence 9, Appl
13	302	47.4	369	12	US-10-194-885-8	Sequence 8, Appl
14	302	47.4	369	12	US-10-091-841-1	Sequence 1, Appl
15	298	46.8	382	12	US-10-091-841-3	Sequence 3, Appl
16	298	46.8	393	12	US-10-091-841-5	Sequence 5, Appl
17	274	43.0	260	10	US-09-878-574-12941	Sequence 12941, A
18	274	43.0	277	10	US-09-878-574-8135	Sequence 8135, Ap
19	263	41.3	272	10	US-09-878-574-12181	Sequence 12181, A
20	258	40.5	402	10	US-09-938-842A-225	Sequence 225, App
21	258	40.5	613	9	US-09-770-149-855	Sequence 855, App
22	257	40.3	268	10	US-09-878-574-5995	Sequence 5995, Ap
23	256	40.2	270	10	US-09-878-574-13445	Sequence 13445, A
24	252	39.6	456	10	US-09-924-035A-677	Sequence 677, App
25	250	39.2	250	9	US-09-923-876-1288	Sequence 1288, Ap
26	245	38.5	370	10	US-09-878-574-3169	Sequence 3169, Ap
27	232	36.4	413	11	US-09-918-995-6767	Sequence 6767, Ap
28	231	36.3	413	14	US-10-102-524-1396	Sequence 1396, Ap
29	231	36.3	418	10	US-09-796-692-3641	Sequence 3641, Ap
30	231	36.3	418	14	US-10-040-862-3641	Sequence 3641, Ap
31	231	36.3	464	10	US-09-998-598-1059	Sequence 1059, Ap
32	231	36.3	479	11	US-09-535-459-1145	Sequence 1145, Ap
33	231	36.3	536	14	US-10-066-543-65	Sequence 65, Appl
34	231	36.3	540	10	US-09-920-300A-1247	Sequence 1247, Ap
35	231	36.3	540	13	US-10-033-528-1247	Sequence 1247, Ap
36	231	36.3	557	10	US-09-884-441-88	Sequence 88, Appl
37	231	36.3	557	11	US-09-907-969-88	Sequence 88, Appl
38	231	36.3	557	14	US-10-198-053-88	Sequence 88, Appl
39	231	36.3	568	11	US-09-535-459-1139	Sequence 1139, Ap
40	231	36.3	570	11	US-09-535-459-1149	Sequence 1149, Ap
41	231	36.3	594	10	US-09-884-441-87	Sequence 87, Appl
42	231	36.3	594	11	US-09-907-969-87	Sequence 87, Appl
43	231	36.3	594	14	US-10-198-053-87	Sequence 87, Appl
44	231	36.3	601	10	US-09-884-441-133	Sequence 133, App
45	231	36.3	601	11	US-09-907-969-133	Sequence 133, App

#### ALIGNMENTS

RESULT 1  
US-09-897-898-10  
Sequence 10, Application US/09897898  
Patent No. US20020037303A1  
GENERAL INFORMATION:  
APPLICANT: DECKERS, HARM M.  
APPLICANT: VAN ROOIJEN, GIJS  
APPLICANT: BOONHE, JOSEPH  
APPLICANT: GOLL, JANSIS  
APPLICANT: MOLONEY, MAURICE M.  
TITLE OF INVENTION: THIOREDOXIN AND THIOREDOXIN REDUCTASE CONTAINING OIL  
FILE REFERENCE: 034547/0104  
CURRENT APPLICATION NUMBER: US/09/897, 898  
CURRENT FILING DATE: 2001-09-21  
PRIOR APPLICATION NUMBER: 09/577,147  
PRIOR FILING DATE: 2000-05-24  
PRIOR APPLICATION NUMBER: 09/448,600  
PRIOR FILING DATE: 1999-11-24  
PRIOR APPLICATION NUMBER: 09/084,777  
PRIOR FILING DATE: 1998-05-27  
PRIOR APPLICATION NUMBER: 60/047,753  
PRIOR FILING DATE: 1997-05-27  
PRIOR APPLICATION NUMBER: 60/047,779

; PRIOR FILING DATE: 1997-05-28  
; PRIOR APPLICATION NUMBER: 60/075,863  
; PRIOR FILING DATE: 1998-02-25  
; PRIOR APPLICATION NUMBER: 60/075,864  
; PRIOR FILING DATE: 1998-02-25  
; NUMBER OF SEQ ID NOS: 24  
; SOFTWARE: Patent Ver. 2.1  
; SEQ ID NO 10  
; LENGTH: 3888  
; TYPE: DNA  
; ORGANISM: Unknown Organism  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: (1555)..(2250)  
; NAME/KEY: CDS  
; LOCATION: (2491)..(2655)  
; OTHER INFORMATION: Description of Unknown Organism: Phaseolin  
; OTHER INFORMATION: promoter-trxn oleosin-phaseolin terminator  
US-09-897-898-10

Alignment Scores:  
Pred. No.: 2,2e-43 Length: 3888  
Score: 377.00 Matches: 73  
Percent Similarity: 71.67% Conservative: 13  
Best Local Similarity: 60.83% Mismatches: 34  
Query Match: 59.18% Indels: 0  
Gaps: 0

US-09-786-715-2 (1-122) x US-09-897-898-10 (1-3888)

QY 2 AAGGCTTCAATCCAAACTCTTGTGCTGATTTCACGGCTTCTTGTGACCA 1620  
DB 1561 TCGGAAGAGAGAGATGATCGCTGCCACCGCTTGAGACATGGAACGAGCTTCAG 1620  
QY 22 LysTrpLysAspSerGluLysLeuValValAlaAspPheThrAlaSerTPCysGlyPro 41  
DB 1621 AAGGCTTCAATCCAAACTCTTGTGCTGATTTCACGGCTTCTTGTGACCA 1680  
QY 42 CysArgAlaIleAlaProTyrPheThrGluLeuAlaLysAsnAsnProAsnValAlaPhe 61  
DB 1681 TGTGCTTTCATCTCATCTTCTTGTGCTGATTTCAGTAAGAACTTCTTACGTGCTTTC 1740  
QY 62 LeuLysValAspValAspGluLeuAsnSerValAlaSerLysTrpGluIleAsnAlaMet 81  
DB 1741 CTCAGCTTCATCTCATCTTGTGCTGATTTCAGTAAGTAATGCGGATGCGGATACAGCGCATG 1800  
QY 82 ProThrPheValPheLeuLysLysGlyLysIleIleGluLysIleValGlyAlaAspLys 101  
DB 1801 CCAACCTTCATGTTTGAAGGAAGGAAGATTTTGGACAAGATTTTGGAGCCAGAAA 1860  
QY 102 ValGlyLeuSerLysLysIleLeuGluLeuSerGlyThrThrProAlaAlaThrSerThr 121  
DB 1861 GATGAGCTTCATCTCATCTTTCAGTAAGAACTTGGCTATGCGGATACAGCTAGAGAAC 1920

RESULT 2  
US-09-938-842A-163  
; Sequence 163, Application US/09938842A  
; Patent No. US20020160378A1  
; GENERAL INFORMATION:  
; APPLICANT: Harper, Jeff  
; APPLICANT: Kreps, Joel  
; APPLICANT: Wang, Xun  
; APPLICANT: Zhu, Tong  
; TITLE OF INVENTION: STRESS-REGULATED GENES OF PLANTS, TRANSGENIC PLANTS CONTAINING  
; FILE REFERENCE: SCRIPI300-3  
; CURRENT APPLICATION NUMBER: US/09/938,842A  
; CURRENT FILING DATE: 2001-08-24  
; PRIOR APPLICATION NUMBER: US 60/227,866  
; PRIOR FILING DATE: 2000-08-24  
; PRIOR APPLICATION NUMBER: US 60/264,647  
; PRIOR FILING DATE: 2001-01-16  
; PRIOR APPLICATION NUMBER: US 60/300,111

; PRIOR FILING DATE: 2001-06-22  
; NUMBER OF SEQ ID NOS: 5379  
; SEQ ID NO 163  
; LENGTH: 345  
; TYPE: DNA  
; ORGANISM: Arabidopsis thaliana  
US-09-938-842A-163

Alignment Scores:  
Pred. No.: 8,69e-45 Length: 345  
Score: 376.00 Matches: 71  
Percent Similarity: 77.57% Conservative: 12  
Best Local Similarity: 66.36% Mismatches: 24  
Query Match: 59.03% Indels: 0  
Gaps: 0

US-09-786-715-2 (1-122) x US-09-938-842A-163 (1-345)

QY 2 AAGGCTTCAATCCAAACTCTTGTGCTGATTTCACGGCTTCTTGTGACCA 126  
DB 7 TCGGAAGAGAGAGATGATCGCTGCCACCGCTTGAGACATGGAACGAGCTTCAG 66  
QY 22 LysTrpLysAspSerGluLysLeuValValAlaAspPheThrAlaSerTPCysGlyPro 41  
DB 67 AAGGCTTCAATCCAAACTCTTGTGCTGATTTCACGGCTTCTTGTGACCA 126  
QY 42 CysArgAlaIleAlaProTyrPheThrGluLeuAlaLysAsnAsnProAsnValAlaPhe 61  
DB 127 TGTGCTTTCATCTCATCTTCTTGTGCTGATTTCAGTAAGAACTTCTTACGTGCTTTC 186  
QY 62 LeuLysValAspValAspGluLeuAsnSerValAlaSerLysTrpGluIleAsnAlaMet 81  
DB 187 CTCAGCTTCATCTCATCTTGTGCTGATTTCAGTAAGTAATGCGGATGCGGATACAGCGCATG 246  
QY 82 ProThrPheValPheLeuLysLysGlyLysIleIleGluLysIleValGlyAlaAspLys 101  
DB 247 CCAACCTTCATGTTTGAAGGAAGGAAGATTTTGGACAAGATTTTGGAGCCAGAAA 306  
QY 102 ValGlyLeuSerLysLysIle 108  
DB 307 GATGAGCTTCATCTCATCTTTCAGTAAGAACTTGGCTATGCGGATACAGCTAGAGAAC 327

RESULT 3  
US-09-897-898-5  
; Sequence 5, Application US/09897898  
; Patent No. US20020037030A1  
; GENERAL INFORMATION:  
; APPLICANT: DECKERS, HARM M.  
; APPLICANT: VAN ROOIJEN, GIJS  
; APPLICANT: BOOTHIE, JOSEPH  
; APPLICANT: GOLL, JANIS  
; APPLICANT: MOLONEY, MAURICE M.  
; APPLICANT: DALMIA, BIPIN K.  
; TITLE OF INVENTION: THIOREDOXIN AND THIOREDOXIN REDUCTASE CONTAINING OIL  
; FILE REFERENCE: 034547/0104  
; CURRENT APPLICATION NUMBER: US/09/897,898  
; CURRENT FILING DATE: 2001-09-21  
; PRIOR APPLICATION NUMBER: 09/577,147  
; PRIOR FILING DATE: 2000-05-24  
; PRIOR APPLICATION NUMBER: 09/448,600  
; PRIOR FILING DATE: 1999-11-24  
; PRIOR APPLICATION NUMBER: 09/084,777  
; PRIOR FILING DATE: 1998-05-27  
; PRIOR APPLICATION NUMBER: 60/047,753  
; PRIOR FILING DATE: 1997-05-27  
; PRIOR APPLICATION NUMBER: 60/047,779  
; PRIOR FILING DATE: 1997-05-28  
; PRIOR APPLICATION NUMBER: 60/075,863  
; PRIOR FILING DATE: 1998-02-25  
; PRIOR APPLICATION NUMBER: 60/075,864  
; PRIOR FILING DATE: 1998-02-25  
; NUMBER OF SEQ ID NOS: 24

```

; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 5
; LENGTH: 3129
; TYPE: DNA
; ORGANISM: Arabidopsis thaliana
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1555)..(1896)
US-09-897-898-5

Alignment Scores:
Pred. No.: 2,23e-43 Length: 3129
Score: 376.00 Matches: 71
Percent Similarity: 77.57% Conservative: 12
Best Local Similarity: 66.36% Mismatches: 24
Query Match: 59.03% Indels: 0
DB: Gaps: 0

US-09-786-715-2 (1-122) x US-09-897-898-5 (1-3129)
QY 2 Aaagluglvglyglvallylalaacyshislyslleaspglutrpgluglylncly 21
    ::::::::::::::::::::::::::::
DB 1561 TCGGAAGAAGCAAGATGATCCCTGCCACACCGTTGACATGGAACGACGCTTCAG 1620
QY 22 LysTrpLysAspSerGluLysLeuValValAlaAspPheThrAlaSerTrpGlyPro 41
    ||| ::::::::::::::::::::::::::::::::::::::::::::::::::::::
DB 1621 AAGGCTAATGAATCCAAACTCTTGtGGGTGATTCACGCTTCTTGtGTGACCA 1680
QY 42 CysArgAlaIleAlaProTyrPheThrGluLeuAlaLysAsnAspProAsnValAlaPhe 61
    ||||| ::::::::::::::::::::
DB 1681 TGTCCTTCATCGCCATTCCTTCTGCTGCTAGAAACTCTTAACGCTGCTTTC 1740
QY 62 LeuLysValAspValAspGluLeuAsnSerValAlaSerLysTrpGluIleAsnAlaMet 81
    ||||| ::::::::::::::::::::
DB 1741 CTCAGAGTTGATGATGATGAATGATGCGGTGGAGAGTGGCGCTACAGCGCATG 1800
QY 82 ProThrPheValPheLeuLysGlyLysIleIleGluLysIleValGlyValAlaAspLys 101
    ||||| ::::::::::::::::::::
DB 1801 CCAACCTTCATCTTTTGAAGGAAGGAAGATTTGGACAAAGTTGTTGGACCAAGAA 1860
QY 102 ValGlyLeuSerLysLysIle 108
    |||
DB 1861 GATGAGCTTCAGTCACTACCATTT 1881

RESULT 4
US-09-897-898-7
; Sequence 7, Application US/09897898
; Patent No. US20020037303A1
; GENERAL INFORMATION:
; APPLICANT: DECKERS, HARM M.
; APPLICANT: VAN ROOIJEN, GIJS
; APPLICANT: BOOTHE, JOSEPH
; APPLICANT: GOLL, JANIS
; APPLICANT: MOLONEY, MAURICE M.
; APPLICANT: DALMAI, BIPIN K.
; TITLE OF INVENTION: THIOREDOXIN AND THIOREDOXIN REDUCTASE CONTAINING OIL
; TITLE OF INVENTION: BODY BASED PRODUCTS
; FILE REFERENCE: 034547/0104
; CURRENT APPLICATION NUMBER: US/09/897,898
; CURRENT FILING DATE: 2001-09-21
; PRIOR APPLICATION NUMBER: 09/577,147
; PRIOR FILING DATE: 2000-05-24
; PRIOR APPLICATION NUMBER: 09/448,600
; PRIOR FILING DATE: 1999-11-24
; PRIOR APPLICATION NUMBER: 09/084,777
; PRIOR FILING DATE: 1998-05-27
; PRIOR APPLICATION NUMBER: 60/047,753
; PRIOR FILING DATE: 1997-05-27
; PRIOR APPLICATION NUMBER: 60/047,779
; PRIOR FILING DATE: 1997-05-28
; PRIOR APPLICATION NUMBER: 60/075,863
; PRIOR FILING DATE: 1998-02-25
; PRIOR APPLICATION NUMBER: 60/075,864
```

```

; PRIOR FILING DATE: 1998-02-25
; NUMBER OF SEQ ID NOS: 24
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 7
; LENGTH: 3888
; TYPE: DNA
; ORGANISM: Unknown Organism
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1555)..(1908)
; NAME/KEY: CDS
; LOCATION: (2149)..(2655)
; OTHER INFORMATION: Description of Unknown Organism: Phaseolin
; OTHER INFORMATION: promoter-oleosin Trxh-phaseolin terminator
US-09-897-898-7

Alignment Scores:
Pred. No.: 3.07e-43 Length: 3888
Score: 376.00 Matches: 71
Percent Similarity: 77.57% Conservative: 12
Best Local Similarity: 66.36% Mismatches: 24
Query Match: 59.03% Indels: 0
DB: Gaps: 0

US-09-786-715-2 (1-122) x US-09-897-898-7 (1-3888)
QY 2 Aaagluglvglyglvallylalaacyshislyslleaspglutrpgluglylncly 21
    ::::::::::::::::::::::::::::::::::::
DB 2320 TCGGAAGAAGCAAGATGATCCCTGCCACACCGTTGACATGGAACGACGCTTCAG 2379
QY 22 LysTrpLysAspSerGluLysLeuValValAlaAspPheThrAlaSerTrpGlyPro 41
    ||| ::::::::::::::::::::::::::::::::::::::::::::::::::::::
DB 2380 AAGGCTAATGAATCCAAACTCTTGtGGGTGATTCACGCTTCTTGtGTGACCA 2439
QY 42 CysArgAlaIleAlaProTyrPheThrGluLeuAlaLysAsnAspProAsnValAlaPhe 61
    ||||| ::::::::::::::::::::
DB 2440 TGTCCTTCATCGCCATTCCTTCTGCTGCTAGAAACTCTTAACGCTGCTTTC 2499
QY 62 LeuLysValAspValAspGluLeuAsnSerValAlaSerLysTrpGluIleAsnAlaMet 81
    ||||| ::::::::::::::::::::
DB 2500 CTCAGAGTTGATGATGATGAATGATGCGGTGGAGAGTGGCGCTACAGCGCATG 2559
QY 82 ProThrPheValPheLeuLysGlyLysIleIleGluLysIleValGlyValAlaAspLys 101
    ||||| ::::::::::::::::::::
DB 2560 CCAACCTTCATCTTTTGAAGGAAGGAAGATTTGGACAAAGTTGTTGGACCAAGAA 2619
QY 102 ValGlyLeuSerLysLysIle 108
    |||
DB 2620 GATGAGCTTCAGTCACTACCATTT 2640

RESULT 5
US-09-878-574-3723
; Sequence 3723, Application US/09878574
; Patent No. US20020110548A1
; GENERAL INFORMATION:
; APPLICANT: BYRUM, Joseph R.
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Thompson, Michael D.
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated with
; TITLE OF INVENTION: Plants
; FILE REFERENCE: 38-21(15401)B
; CURRENT APPLICATION NUMBER: US/09/878,574
; CURRENT FILING DATE: 2001-12-21
; PRIOR APPLICATION NUMBER: 09/533,535
; PRIOR FILING DATE: 1999-06-14
; NUMBER OF SEQ ID NOS: 15775
; SEQ ID NO 3723
; LENGTH: 402
; TYPE: DNA
; ORGANISM: Glycine max
; OTHER INFORMATION: Clone ID: LIB3028-006-01-B1-F8
US-09-878-574-3723
```



```
DB      247 ATCCACACCTTGTGTTCATTAAGCCGGGCAAGTTGGATAGCTGTTGGCAAT 306
QY      101 LysValGlyLeuSerLysLysLleuGluLeuSerGlyThrProAla 117
DB      307 AAAGAAAGATCTTCAGCGCAAAATAGTGAAGCATACTGTTGTAACAACGCG 357

RESULT 8
US-09-878-574-1462
; Sequence 1462, Application US/09878574
; Patent No. US20020110548A1
; GENERAL INFORMATION:
; APPLICANT: Byrum, Joseph R.
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Thompson, Michael D.
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated with
; TITLE OF INVENTION: Plants
; FILE REFERENCE: 38-21(15401)B
; CURRENT APPLICATION NUMBER: US/09/878,574
; CURRENT FILING DATE: 2001-12-21
; PRIOR APPLICATION NUMBER: 09/333,535
; PRIOR FILING DATE: 1999-06-14
; NUMBER OF SEQ ID NOS: 15775
; SEQ ID NO 1462
; LENGTH: 390
; TYPE: DNA
; ORGANISM: Glycine max
; OTHER INFORMATION: Clone ID: LIB3028-037-01-B1-F11
US-09-878-574-1462

Alignment Scores:
Pred. No.:      5,866-40      Length:      390
Score:          343.00      Matches:      64
Percent Similarity: 74.53%      Conservative: 15
Best Local Similarity: 60.38%      Mismatches: 27
Query Match:    53.85%      Indels:      0
DB:             10      Gaps:        0

US-09-786-715-2 (1-122) x US-09-878-574-1462 (1-390)
QY      3 GluGluGlyGlnValIleAlaCysHisLysLleAspGluTrpGluGlnLeuGlyLys 22
DB      69 GAAGAGGAGACAGTATCGCGCTGCACACCGTTGATGATGAGTGCACACTGCAGAAAT 128
QY      23 TrpLysAspSerGluLysLeuValValAlaAspPheThrAlaSerTrpCysGlyProCys 42
DB      129 GCAAAAGACTCCAAAAAAGTATGTGTGGTATTTACGCTTCTCTCGTGTGTCATGC 188
QY      43 ArgAlaIleAlaProTyrPheThrGluLeuAlaLysAsnAsnProAsnValAlaPheLeu 62
DB      189 CGTTTATGCGCCCGACGTTCTTCAGAGATTGCCAAAGAAACCTCGAATTGATCTCTC 248
QY      63 LysValAspValAspGluLeuAsnSerValAlaSerLysTrpGluLleAsnAlaMetPro 82
DB      249 AAAGTGATGTGATGAAGTAGGCGCTGTGCTGAGAAATATTCATTGAGCGCATGCCA 308
QY      83 ThrPheValPheLeuLysGlyLysLleuGluLysLleuValGluLysLleuValAlaAspLysVal 102
DB      309 ACCTTCCTCTTCTTTAAAGATGCGCAGATGCTGACACAAAGTGCTGTGCTAGTAAGCAT 368
QY      103 GlyLeuSerLysLysLle 108
DB      369 GACCTTCACAGCACCATTA 386

RESULT 9
US-10-349-782-12
; Sequence 12, Application US/10349782
; Publication No. US20030143618A1
; GENERAL INFORMATION:
; APPLICANT: Yves Hatzfield
; APPLICANT: Valerie Marie-No. US20030143618A1lle Frankard
; APPLICANT: Anne-Marie Droual
; TITLE OF INVENTION: Method for easy cloning and selection of chimeric DNA molecules
; FILE REFERENCE: 1187-15
```

```
; CURRENT APPLICATION NUMBER: US/10/349,782
; CURRENT FILING DATE: 2003-01-23
; PRIOR APPLICATION NUMBER: EP 02075373.7
; PRIOR FILING DATE: 2002-01-23
; NUMBER OF SEQ ID NOS: 12
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 12
; LENGTH: 540
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Modified thioredoxin of Oryza sativa in vector pDONR201
US-10-349-782-12

Alignment Scores:
Pred. No.:      9,466-40      Length:      540
Score:          343.00      Matches:      66
Percent Similarity: 69.75%      Conservative: 17
Best Local Similarity: 55.46%      Mismatches: 36
Query Match:    53.85%      Indels:      0
DB:             12      Gaps:        0

US-09-786-715-2 (1-122) x US-10-349-782-12 (1-540)
QY      2 AlaGluGlyGlnValIleAlaCysHisLysLleAspGluTrpGluGlnLeuGlyLys 21
DB      64 GCCAGAGAGAGGAGTGTGATCCCTGCCCAACAAGACGAGATTCGACGCCCGAGATGACC 123
QY      22 LysTrpLysAspSerGluLysLeuValValAlaAspPheThrAlaSerTrpCysGlyPro 41
DB      124 AAGGCCAAGAGAGCGCGCAAAAGTGTGATTAATGACTTACGCTTCTCGTGGCGGACCG 183
QY      42 CysArgAlaIleAlaProTyrPheThrGluLeuAlaLysAsnAsnProAsnValAlaPhe 61
DB      184 TCCCGCTTCATCGCCCGACGTTCCCTGATACGCCCAAAAGTCCCTGCTGCTCTTC 243
QY      62 LeuLysValAspValAspGluLeuAsnSerValAlaSerLysTrpGluLleAsnAlaMet 81
DB      244 CTGAAGCTTGATGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 303
QY      82 ProThrPheValPheLeuLysLysLleuGluLysLleuGluLysLleuValAlaAspLys 101
DB      304 CCGACCTTCCTATTCATCAAGATGATGCTGAGGCTGACAAAGTGTGCTGCGCCAGAG 363
QY      102 ValGlyLeuSerLysLysLleuGluLeuSerGlyThrThrProAlaIleAsnSer 120
DB      364 GATGACCTTCAGAACACCATGCTGAAGCAGCTGCGTCCATCTGCTCTCT 420

RESULT 10
US-09-938-842A-1736
; Sequence 1736, Application US/09938842A
; Patent No. US20020160378A1
; GENERAL INFORMATION:
; APPLICANT: Harper, Jeff
; APPLICANT: Krepes, Joel
; APPLICANT: Wang, Xun
; APPLICANT: Zhu Tong
; TITLE OF INVENTION: STRESS-REGULATED GENES OF PLANTS, TRANSGENIC PLANTS CONTAINING
; TITLE OF INVENTION: SAME, AND METHODS OF USE
; FILE REFERENCE: SCRIPI300-3
; CURRENT APPLICATION NUMBER: US/09/938,842A
; CURRENT FILING DATE: 2001-08-24
; PRIOR APPLICATION NUMBER: US 60/227,866
; PRIOR FILING DATE: 2000-08-24
; PRIOR APPLICATION NUMBER: US 60/264,647
; PRIOR FILING DATE: 2001-01-16
; PRIOR APPLICATION NUMBER: US 60/300,111
; PRIOR FILING DATE: 2001-06-22
; NUMBER OF SEQ ID NOS: 5379
; SEQ ID NO 1736
; LENGTH: 357
; TYPE: DNA
; ORGANISM: Arabidopsis thaliana
```



```
OY 62 LeuLysValAspValAspGluLeuAsnSerValAlaSerLysTrpGluLeuAsnAlaMet 81
DB 316 CTGAAGTTGATGTTGATGAGCTGAAGAGAGTTGCTGAAAGATACAAATGACAGGCAATG 375
OY 82 ProthrPheValPheLeuLysLysGlyLysIleIleGluLysIleValGlyAlaAspLys 101
DB 376 CCGACCTTCCTCAAT-ATCAAGAGATGCTGAGCTGACCAAGGCTGTTGGCCCAAGAG 434
OY 102 ValGlyLeuSerLysLysIleLeuGluLeuSerGlyThrThrProAlaAlaThrSer 120
DB 435 GATGACCTCCAGACACCATCGTGAACGCTGGGCGCCATCGCTCACTCTCTCT 491

RESULT 13
US-10-194-885-8
; Sequence 8, Application US/10194885
; Publication No. US20030135678A1
; GENERAL INFORMATION:
; APPLICANT: Wong, J. H.
; APPLICANT: Cho, Myeong-Je
; APPLICANT: Lemaux, Peggy G.
; APPLICANT: Buchanan, Bob
; TITLE OF INVENTION: TRANSGENIC PLANTS WITH ELEVATED
; TITLE OF INVENTION: THIODEDOLIN LEVELS
; FILE REFERENCE: 41627200800
; CURRENT FILING DATE: 2002-07-12
; PRIOR FILING DATE: 2001-07-19
; PRIOR APPLICATION NUMBER: 60/307,006
; PRIOR FILING DATE: 2001-07-19
; PRIOR APPLICATION NUMBER: 09/538,864
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: 60/126,736
; PRIOR FILING DATE: 1999-03-29
; NUMBER OF SEQ ID NOS: 55
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 8
; LENGTH: 369
; TYPE: DNA
; ORGANISM: Barley
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)...(369)
US-10-194-885-8

Alignment Scores:
Pred. No.: 4,31e-34 Length: 369
Score: 302.00 Matches: 53
Percent Similarity: 72.90% Conservative: 25
Best Local Similarity: 49.53% Mismatches: 29
Query Match: 47.41% Indels: 0
DB: 12 Gaps: 0

US-09-786-715-2 (1-122) x US-10-194-885-8 (1-369)
OY 2 AlaGluGluGlyGlnValIleAlaCysHisLysIleAspGluTrpGluGlyLeuGly 21
DB 25 GCAGTGGCGGGCGAGGTGATCTCGGTCACACCGCTGGAGCACTGAGATCGAG 84
OY 22 LysTrpLysAspSerGluLysLeuValValAlaAspPheThrAlaSerTrpCysGlyPro 41
DB 85 GAGGCCAACACCGCCAGAGAGTGGTGGATGACTTCACATGCATCAAGTGGCGAGCA 144
OY 42 CysArgAlaIleAlaProTyrPheThrGluLeuAlaLysAsnAsnProAsnValAlaPhe 61
DB 145 TCCCCCATCATGCTCCAGTTTCCGCTGATCTCGCCAGAAAGTCCCAATGCTGTTTC 204
OY 62 LeuLysValAspValAspGluLeuAsnSerValAlaSerLysTrpGluLeuAsnAlaMet 81
DB 205 CTCAGAGTCGAGCTGATGAAGTGAAGCCCATGCTGAGCATTCAGTGTGAGGCCATG 264
OY 82 ProthrPheValPheLeuLysLysGlyLysIleIleGluLysIleValGlyAlaAspLys 101
DB 265 CCAAGCTTCCTTCATGAAGAGAGACGTCAGAGAGGCTTGTGAGACTTATCAAG 324
```

```
OY 102 ValGlyLeuSerLysLysIle 108
DB 325 GAGGAAGTACCCGCAAGGTT 345

RESULT 14
US-10-091-841-1
; Sequence 1, Application US/10091841
; Publication No. US20030150010A1
; GENERAL INFORMATION:
; APPLICANT: Cho, Myeong-Je
; APPLICANT: Del Val, Greg
; APPLICANT: Caillau, Maxime
; APPLICANT: Lemaux, Peggy G.
; APPLICANT: Buchanan, Bob B.
; TITLE OF INVENTION: Barley Gene for Thioredoxin and
; TITLE OF INVENTION: NADP-Thioredoxin Reductase
; FILE REFERENCE: 2001-0701.30
; CURRENT FILING DATE: 2002-03-05
; PRIOR FILING DATE: 09/540,014
; PRIOR FILING DATE: 2000-03-31
; PRIOR APPLICATION NUMBER: US 60/127,198
; PRIOR FILING DATE: 1999-03-31
; PRIOR APPLICATION NUMBER: US 60/169,162
; PRIOR FILING DATE: 1999-12-06
; PRIOR APPLICATION NUMBER: US 60/177,740
; PRIOR FILING DATE: 2000-01-21
; PRIOR APPLICATION NUMBER: US 60/177,739
; PRIOR FILING DATE: 2000-01-21
; NUMBER OF SEQ ID NOS: 51
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 369
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: barley thioredoxin h cDNA
US-10-091-841-1

Alignment Scores:
Pred. No.: 4,31e-34 Length: 369
Score: 302.00 Matches: 53
Percent Similarity: 72.90% Conservative: 25
Best Local Similarity: 49.53% Mismatches: 29
Query Match: 47.41% Indels: 0
DB: 12 Gaps: 0

US-09-786-715-2 (1-122) x US-10-091-841-1 (1-369)
OY 2 AlaGluGluGlyGlnValIleAlaCysHisLysIleAspGluTrpGluGlyLeuGly 21
DB 25 GCAGTGGCGGGCGAGGTGATCTCGGTCACACCGCTGGAGCACTGAGATCGAG 84
OY 22 LysTrpLysAspSerGluLysLeuValValAlaAspPheThrAlaSerTrpCysGlyPro 41
DB 85 GAGGCCAACACCGCCAGAGAGTGGTGGATGACTTCACATGCATCAAGTGGCGAGCA 144
OY 42 CysArgAlaIleAlaProTyrPheThrGluLeuAlaLysAsnAsnProAsnValAlaPhe 61
DB 145 TCCCCCATCATGCTCCAGTTTCCGCTGATCTCGCCAGAAAGTCCCAATGCTGTTTC 204
OY 62 LeuLysValAspValAspGluLeuAsnSerValAlaSerLysTrpGluLeuAsnAlaMet 81
DB 205 CTCAGAGTCGAGCTGATGAAGTGAAGCCCATGCTGAGCATTCAGTGTGAGGCCATG 264
OY 82 ProthrPheValPheLeuLysLysGlyLysIleIleGluLysIleValGlyAlaAspLys 101
DB 265 CCAAGCTTCCTTCATGAAGAGAGACGTCAGAGAGGCTTGTGAGACTTATCAAG 324
```

```
RESULT 15
US-10-091-841-3
; Sequence 3, Application US/10091841
; Publication No. US20030150010A1
; GENERAL INFORMATION:
; APPLICANT: Cho, Myeong-Je
; APPLICANT: Del Val, Greg
; APPLICANT: Calliau, Maxime
; APPLICANT: Lemauz, Peggy G.
; APPLICANT: Buchanan, Bob B.
; TITLE OF INVENTION: Barley Gene for Thioredoxin and
; FILE REFERENCE: 2001-0701.30
; CURRENT APPLICATION NUMBER: US/10/091,841
; PRIOR APPLICATION NUMBER: 09/540,014
; PRIOR FILING DATE: 2000-03-31
; PRIOR APPLICATION NUMBER: US 60/127,198
; PRIOR FILING DATE: 1999-03-31
; PRIOR APPLICATION NUMBER: US 60/169,162
; PRIOR FILING DATE: 1999-12-06
; PRIOR APPLICATION NUMBER: US 60/177,740
; PRIOR FILING DATE: 2000-01-21
; PRIOR APPLICATION NUMBER: US 60/177,739
; PRIOR FILING DATE: 2000-01-21
; NUMBER OF SEQ ID NOS: 51
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 3
; LENGTH: 382
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: wheat thioredoxin h cDNA
US-10-091-841-3

Alignment Scores:
Pred. No.: 1,71e-33 Length: 382
Score: 298.00 Matches: 52
Percent Similarity: 74.04% Conservative: 25
Best Local Similarity: 50.00% Mismatches: 27
Query Match: 46.78% Indels: 0
DB: 12 Gaps: 0

US-09-786-715-2 (1-122) x US-10-091-841-3 (1-382)
QY 5 GlyValIleAlaCysHisLysIleAspGluTrpGluGlnLeuGlyLysTrpLys 24
   |||:::||||::: ||| :::::||||| |||::: ::::
Db 52 GGGGAGGTGATCTCCGCCACAGCTGAGCATGCAGATCGAGAGGCCAAC 111
QY 25 AspSerGlnLysLeuValValIleAspPheThrAlaSerTrpCysGlyProCysArgAla 44
   :::::|||||:::|||||:::|||||:::|||||:::|||||:::
Db 112 GCCGCCAAGAGCTGGTGGATGACTTCATCATCATGATGTCGCGACCATGCCGCATT 171
QY 45 IleAlaProThrPheThrGlnLeuAlaLysAsnAsnProAsnValAlaPheLeuLysVal 64
   :::::||||| ||| :::::||||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 172 ATGGCTCCAAATTTCGCTGATCTGCCAAGAGCTCCACAGCTGCTGTTTCTCAAGGTC 231
QY 65 AspValAspGlnLeuAsnSerValAlaSerLysTrpGlnIleAsnAlaMetProThrPhe 84
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 232 GAGCTGATGATCAAGAGCCATTGCTGAGCAATTCAGCTGGAGGCCATGCCACCTTC 291
QY 85 ValPheLeuLysLysGlyLysIleIleGlnLysIleValGlyAlaAspLysValGlyLeu 104
   :::::|||||::: ||| :::::|||||::: ||| ||| ||| ||| ||| ||| ||| |||
Db 292 CTGTCATGAAAGAGATGTCACAGACAGGCTGTGCGACCTATCAAGAGGAACCTG 351
QY 105 SerLysLysIle 108
   :::: |||:::
Db 352 ACGACCAAGGTT 363
```

Search completed: August 17, 2003, 23:43:46  
Job time: 146.207 secs



GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - nucleic search, using frame\_plus\_p2n model

Run on: August 17, 2003, 19:20:27 ; Search time 1649.9 Seconds  
(without alignments)  
1797.162 Million cell updates/sec

Title: US-09-786-715-2  
Perfect score: 637  
Sequence: 1 MAEEOVYACHKIDWEGQL.....GLSKKLELSTTPATSTA 122

Scoring table:  
BLOSUM62  
Xgapop 10.0 , Xgapext 0.5  
Ygapop 10.0 , Ygapext 0.5  
Fgapop 6.0 , Fgapext 7.0  
Delop 6.0 , Delext 7.0

Searched: 22781392 segs, 12152238056 residues  
Total number of hits satisfying chosen parameters: 45562784

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Command line parameters:  
-MODEL=frame+2n model -DEV=xlp  
-O=/ggn2.1/USPTO/US09786715/runat\_11082003\_150514\_6048/app.query.fasta\_1.1052  
-DB=EST -QFMT=fastap -SUFFIX=first -MINMATCH=0.1 -LOOPEL=0 -LOOPEXT=0  
-ONITS=bits -STAR=1 -END=-1 -MATRIX=biosum62 -TRANS=human4.0.cdi -LIST=45  
-DOCLIGN=200 -THR.SCORE=Pct -THR.MAX=100 -THR.MIN=0 -ALIGN=15 -MODE=LOCAL  
-OUTFMT=nto -NORM=ext -HEADSIZE=500 -MINLEN=0 -MAXLEN=200000000  
-USER=US09786715\_@CGN\_1.1.7834\_@runat\_11082003\_150514\_6048 -NCPU=6 -ICPU=3  
-NO\_MMAP -LARGEQUERY -NEG.SCORES=0 -WAIT -DSPLOCK=100 -LONGLOC  
-DEV.TIMEOUT=120 -WARN.TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FCGAPOP=6  
-XGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :  
EST:\*  
1: em\_estba:\*  
2: em\_estcha:\*  
3: em\_estlin:\*  
4: em\_estmu:\*  
5: em\_estrov:\*  
6: em\_estrpl:\*  
7: em\_estro:\*  
8: em\_hlc:\*  
9: gb\_est1:\*  
10: gb\_est2:\*  
11: gb\_hrc:\*  
12: gb\_est3:\*  
13: gb\_est4:\*  
14: gb\_est5:\*  
15: em\_estfun:\*  
16: em\_estom:\*  
17: em\_gss\_hum:\*  
18: em\_gss\_inv:\*  
19: em\_gss\_pln:\*  
20: em\_gss\_fyt:\*  
21: em\_gss\_yut:\*  
22: em\_gss\_mam:\*  
23: em\_gss\_mam:\*  
24: em\_gss\_pro:\*  
25: em\_gss\_tod:\*  
26: em\_gss\_phg:\*  
27: em\_gss\_vr1:\*  
28: gb\_gss1:\*

29: gb\_gss2:.\*  
Pred.. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	390.5	61.3	622	9	AW255457
2	385	60.4	487	13	B0878216
3	385	60.4	538	9	A1161830
4	385	60.4	568	13	B0866690
5	385	60.4	610	13	B0836847
6	385	60.4	611	14	CA821710
7	385	60.4	612	13	B0888094
8	385	60.4	615	13	B0830685
9	385	60.4	641	13	B0873533
10	383	60.1	520	13	B0885783
11	383	60.1	523	13	B0812532
12	383	60.1	523	13	B0829684
13	383	60.1	531	12	B1136756
14	383	60.1	532	13	B0884977
15	383	60.1	544	13	B0827985
16	383	60.1	555	13	B0885050
17	383	60.1	557	13	B0886021
18	383	60.1	579	13	B0876591
19	383	60.1	591	13	B0884811
20	383	60.1	594	13	B0837287
21	383	60.1	600	13	B0884129
22	383	60.1	601	13	B0895406
23	383	60.1	670	13	B0868860
24	382.5	60.0	602	10	BE054543
25	382	60.0	492	12	B1126457
26	382	60.0	495	12	B1124724
27	382	60.0	507	12	B1124769
28	382	60.0	534	12	B1124332
29	382	60.0	560	12	B1126407
30	382	60.0	611	13	B0872611
31	381	59.8	591	14	CA823905
32	381	59.8	591	14	CA826440
33	379	59.5	397	13	B0816912
34	379	59.5	461	12	B1125905
35	379	59.5	467	13	B0818910
36	379	59.5	485	13	B0818945
37	379	59.5	570	12	B1126033
38	379	59.5	588	14	CB185215
39	379	59.5	603	13	B0874273
40	379	59.5	606	13	B0408049
41	379	59.5	622	14	CB239393
42	379	59.5	658	13	B0874439
43	378.5	59.4	455	9	A1773303
44	378.5	59.4	477	9	AM041395
45	378.5	59.4	505	9	AM094525

## ALIGNMENTS

RESULT 1  
AW255457  
LOCUS M480 peppermint glandular trichome Mentha x piperita CDNA, mRNA  
DEFINITION sequence.  
ACCESSION AW255457  
VERSION AW255457.1  
KEYWORDS GI:7244709  
SOURCE Menth x piperita (peppermint)  
ORGANISM Menth x piperita  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
asterids; lamiales; Lamiaceae; Nepetoideae; Menthae;

REFERENCE 1 (bases 1 to 622)  
 AUTHORS Lange, B.M., Wildung, M.R., Stauber, E.J., Sanchez, C., Pouchnik, D. and Croteau, R.  
 TITLE Probing essential oil biosynthesis and secretion by functional evaluation of expressed sequence tags from mint glandular trichomes  
 JOURNAL Proc. Natl. Acad. Sci. U.S.A. 97 (6), 2934-2939 (2000)  
 MEDLINE 20183992  
 PUBMED 10717007  
 COMMENT Contact: Lange, B.M.  
 Institute of Biological Chemistry/Washington State University  
 Pullman, WA  
 Email: lange-memal.wsu.edu.  
 Location/Qualifiers  
 FEATURES  
 source 1..622  
 /organism="Mentha x piperita"  
 /mol\_type="mRNA"  
 /cultiyar="Black Mitcham"  
 /db\_xref="taxon:34256"  
 /lssue\_type="petalate glandular trichomes"  
 /cell\_type="secretory"  
 /clone\_idb="peppermint glandular trichome"  
 /note="Vector: lambda ZAPII"  
 BASE COUNT 192 a 129 c 143 g 158 t  
 ORIGIN  
 Alignment Scores:  
 Pred. No.: 7,16e-40 Length: 622  
 Score: 390.50 Matches: 77  
 Percent Similarity: 73.95% Conservative: 11  
 Best Local Similarity: 64.71% Mismatches: 28  
 Query Match: 61.30% Indels: 3  
 DB: Gaps: 1  
 US-09-786-715-2 (1-122) x AM255457 (1-622)  
 QY 4 GUGUGYUValIIleAlaCySHISLysLIleAspGUtrpGUInleuGlyLysTrp 23  
 Db 65 GAAGGACAGGTGATGGCTGCCACACACATGACCTGGAACGAGCTTCAGAAAGCG 124  
 QY 24 LysAspSerGIuLysLeuValIValAspPheThrAlaSerTPCySGlyProCyArg 43  
 Db 125 NAGGATACAGAAAGGTGAGTAGTGTGATTTCTCTCTCTCTGTCGAGACCTGTCGG 184  
 QY 44 AlaIleAProfYrPheThrGIuLeuAlaLysAsnAsnProAsnValAlaPheLeuLys 63  
 Db 185 TTGATCGCTCCATCTTCGCGAAGATTGCGCAAGAGTTCCCTATGTGACATTTCTCAAG 244  
 QY 64 ValAspValAspGIuLeuAsnSerValAlaSerLysTrpGUleuAsnAlaMetProThr 83  
 Db 245 GTGGAGTGTGATGAGTGAAGCTGGTGTGACTGAGCGAGCGAGCGCAATGCCAAC 304  
 QY 84 PheValPheLeuLysLysGIYLSIleIleGIuLysLIleValAlaGlyAlaAspLysValGly 103  
 Db 305 TTTCATCTTCTCTCAAGAAAGGAAATTTTGACAGATCTCTGGAGCGAAAGAAAGAG 364  
 QY 104 LeuSerLysLysLIleLeuGIuLeuSerGIYThrThProAlaIleThrSerThrAla 122  
 Db 365 CTCGAGCGCAATATTTGCTAGACACCTCAACACA-----GCTACTACTACTGCT 412  
 RESULT 2  
 LOCUS BU878216 487 bp mRNA linear EST 16-OCT-2002  
 DEFINITION trichocarpa cDNA 5 prime, mRNA sequence.  
 ACCESSION BU878216  
 VERSION BU878216.1 GI:24069740  
 SOURCE EST.  
 ORGANISM Populus balsamifera subsp. trichocarpa  
 Populus balsamifera subsp. trichocarpa  
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids  
 ; eustosids I; Malpighiales; Salicaceae; Populus.

REFERENCE 1 (bases 1 to 487)  
 AUTHORS Umeberg, P., Bhalerao, R.R., Jansson, S. and Sterky, F.  
 TITLE The poplar tree transcriptome: Analysis of expressed sequence tags from multiple libraries  
 JOURNAL Unpublished  
 COMMENT Contact: BHALERAO RUPALI R.  
 Umea Plant Science Center  
 Department of Plant Physiology  
 University of Umea, 901 87 Umea, Sweden  
 Tel: +46 90 786 5279  
 Fax: +46 90 786 6676  
 Email: rupali.bhalerao@plantphys.umu.se.  
 Location/Qualifiers  
 FEATURES  
 source 1..487  
 /organism="Populus balsamifera subsp. trichocarpa"  
 /mol\_type="mRNA"  
 /sub\_species="trichocarpa"  
 /db\_xref="taxon:3694"  
 /clone\_idb="Populus flower cDNA library"  
 /note="Organ: flower"  
 BASE COUNT 130 a 80 c 127 g 150 t  
 ORIGIN  
 Alignment Scores:  
 Pred. No.: 2.63e-39 Length: 487  
 Score: 385.00 Matches: 70  
 Percent Similarity: 78.85% Conservative: 12  
 Best Local Similarity: 67.31% Mismatches: 22  
 Query Match: 60.44% Indels: 0  
 DB: Gaps: 13  
 US-09-786-715-2 (1-122) x BU878216 (1-487)  
 QY 1 MetValaGIuGIuGIuValIIleAlaCySHISLysLIleAspGUtrpGUInleu 20  
 Db 33 ATGGCCGAAAGAGACAGTATTGCTGCCACACAGTATCTGGAAGACATTTC 92  
 QY 21 GYLYSTrPysAspSerGIuLysLeuValIValAspPheThrAlaSerTPCySGly 40  
 Db 93 GAGAGGGAAAGAGGCTCTCAAGAACTGATGCTGTGATTTACTCTCTCATGCTGTC 152  
 QY 41 ProCysArgAlaIleAProfYrPheThrGIuLeuAlaLysAsnAsnProAsnValAla 60  
 Db 153 CCATGTAAATGATGTCTCAATCTTCGCGAGTGGCGAAGAGTTCCCAATGTCA 212  
 QY 61 PheLeuLysValAspValAspGIuLeuAsnSerValAlaSerLysTrpGIuIleAsnAla 80  
 Db 213 TTCTTGAAGTGTGATGTGATGATTAAGCTGTCTGAGAGTGTGATGTGAGGCA 272  
 QY 81 MetCProThrPheValPheLeuLysLysGIYLSIleIleGIuLysLIleValAlaAsp 100  
 Db 273 ATGCCAACTTTATTTCCTGAAAGATGAAATTTAGTGACAAACTGTGGGTGCTGAT 332  
 QY 101 LysValGIuLeu 104  
 Db 333 AAGATGCGCTG 344  
 RESULT 3  
 LOCUS A1161830 538 bp mRNA linear EST 03-DEC-1998  
 DEFINITION A007P520 Hybrid aspen plasmid library Populus tremula x Populus tremuloides cDNA 5', mRNA sequence.  
 ACCESSION A1161830  
 VERSION A1161830.1 GI:3853115  
 SOURCE EST.  
 ORGANISM Populus tremula x Populus tremuloides  
 Populus tremula x Populus tremuloides  
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids  
 ; eustosids I; Malpighiales; Salicaceae; Populus.  
 REFERENCE 1 (bases 1 to 538)  
 Sterky, F., Regan, S., Karlsson, J., Hertzberg, M., Rohde, A., Holmberg  
 A., Amin, B., Bhalerao, R., Larsson, M., Villarroel, R., Van Montagu



ACCESSION	B0836847
VERSION	B0836847.1
KEYWORDS	GI:24019659
SOURCE	EST.
ORGANISM	Populus tremula x Populus tremuloides Populus tremula x Populus tremuloides Eukaryota; Viridiplantae; Streptophyta; Eumariophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids ; eurosoids I; Malpighiales; Salicaceae; Populus.
REFERENCE	1 (bases 1 to 610)
AUTHORS	Umeberg,P., Bhalerao,R.R., Jansson,S. and Sterky,F.
TITLE	The poplar tree transcriptome: Analysis of expressed sequence tags from multiple libraries
JOURNAL	Unpublished
COMMENT	Contact: BHALERAO RUPALI R. Umea Plant Science Center Department of Plant Physiology University of Umea, 901 87 Umea, Sweden Tel.: +46 90 786 5279 Fax: +46 90 786 6676 Email: rupali.bhaleraoplantphys.umu.se.
FEATURES	Location/Qualifiers
source	1..610
	/organism="Populus tremula x Populus tremuloides"
	/mol_type="mRNA"
	/db_xref="taxon:47664"
	/tissue_type="apical shoot"
	/clone_lib="Populus apical shoot cDNA library"
BASE COUNT	153 a 95 c 160 g 202 t
ORIGIN	
Alignment Scores:	
Pred. No.:	3.57e-39 Length: 610
Score:	385.00 Matches: 70
Percent Similarity:	78.85% Conservative: 12
Best Local Similarity:	67.31% Mismatches: 22
Query Match:	60.44% Indels: 0
DB:	13 Gaps: 0
US-09-786-715-2 (1-122) x B0836847 (1-610)	
Oy	1 MetAAGlUGlUGlGAlValLleAlacysHISlysIleAspgLUTrPGlUGlYnLeu 20 
Dd	34 ATGCCCGAAGAAGCACAAGTTATTCCTCCGCACACAGTGGATTACC TGAAAGACATTTC 93 
Oy	21 GLVlySTRPlYsASPserGLULysLeuValValaSPherThraLaSerTrPCysGLy 40 
Dd	94 GAGAAAGGAAGAGGGGTCTCAGAAACTGATGCGTGATTTTACTGCTTCATGGGTCCA 153 
Oy	41 PRCYSatgAlaIleAlaPrtOryrPherThrgLUeuAlaLySaSNasPrCaSNVala 60 
Dd	154 CCAGTATAAAAGATGCTCCCACTCTCCGCCAGTTGGCAGAAAGTTCTTAATGTCCACA 213 
Oy	61 PhelEuLYsVALaspVALaspCLULEuanserValAlaSerLYstrPGLIILeASnla 80 
Dd	214 TTCTTGAAAGGTGAAGTGAATGAAGGCTGTGTCTGAAGAGTGAATGTGAAGGCA 273 
Oy	81 MeTPRThrPheValPhelEuLYsGLysGLysILElleGLULysILEVaIGlYALaasp 100 
Dd	274 ATGCCAACATTATTTTCTCAAGAGATGCAAAAATTAGTGCACAAACATGTGGTCTCAT 333 
Oy	101 LysValGLyLeu 104 
Dd	334 AAAGATGATCTCG 345
RESULT 6	
LOCUS	CA821710 611 bp mRNA linear EST 28-FEB-2003
DEFINITION	RS406D10 two-month-old roots from clone 'Beaufre' grown for 19 days under restricted irrigation Populus balsamifera subsp. trichocarpa
ACCESSION	CA821710
VERSION	CA821710.1
KEYWORDS	GI:28605259
EST.	

SOURCE ORGANISM	Populus balsamifera subsp. trichocarpa x Populus deltoides Populus balsamifera subsp. trichocarpa x Populus deltoides Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophytes; Magnoliophyta; eudicotyledons; core eudicots; rosids ; eurosid1; Malpighiales; Salicaceae; Populus.
REFERENCE AUTHORS	1 (bases 1 to 611) Kohler,A., Delaruelle,C., Martin,D. and Martin,F.
TITLE	The poplar root transcriptome: analysis of 6000 expressed sequence tags
JOURNAL COMMENT	Unpublished (2002) Contact: Martin FM Equipe de Microbiologie Forestiere Institut National de la Recherche Agronomique Centre INRA de Nancy, 54280 Champenoux, France Tel.: +33 383 39 40 80 Fax: +33 383 39 40 69 Email: fmartin@nancy.inra.fr Insert length: 611 std Error: 0.00 Seq primer: Fornat 5' AAGCGCCGATTTGTGTGTCACC. Location/Qualifiers
FEATURES source	1..611 /organism="Populus balsamifera subsp. trichocarpa x Populus deltoides" /mol_type="mRNA" /cultivar="Beaupre/" /db_xref="taxon:3695" /dev_stage="two-month-old" /clone_lib="two-month-old roots from clone 'Beaupre' grown for for 19 days under restricted irrigation" /note="Organ: root; Vector: pTriblex2; cDNA library of roots from two-month-old Populus trichocarpa Torr & Gray x deltoides Bartr. Ex Marshall (clone 'Beaupre') grown for 19 days under restricted irrigation to reach 50% of the transpiration rate of fully watered plants. The cDNA library was constructed from 1 ug of total RNA using the SMART cDNA synthesis kit (Clontech, Palo Alto, CA, USA) according to the manufacturer's instructions. The resulting cDNA was packed into lambda phages using the Gispack III Gold packaging kit (Stratagene, La Jolla, CA . The pTriblex2 phagemid clones in Escherichia coli obtained by using the mass in vivo excision protocol . according to the manufacturer's instructions (Clontech)."
BASE COUNT	161 a 97 c 159 g 190 t 4 others
ORIGIN	
Alignment Scores:	
Pred. No.:	3.58e-39 Length: 611
Score:	385.00 Matches: 70
Percent Similarity:	78.85% Conservative: 12
Best local Similarity:	67.31% Mismatches: 22
Query Match:	60.44% Indels: 0
DB:	14 Gaps: 0
OS-09-786-715-2 (1-122) x CA821710 (1-611)	
OY 1	MetlaagiugluglgiyginValilleAlAcyshtisylsileaspgiutrgplugiLndeu 20      .....    :       ::: Db 58 ATGGCCGAAGAAGACCAAGTAATTGCCGCACACAGTGACTGTGAAGAAGACATTTC 117 21 GlytistrplysaspsercLiLysleuVaIValaspPheThrAlAsertPCysgly 40         .....    :        :        Db 118 GAGAAAGGAAAAGGCTGCAGAAACGTGATTCGTGCTGATTACTGCTTCATGCTGCA 177 OY 41 ProCysargAlAlleAlaprotYrPhenhrcLndeuAlalsasnaAnProasValAla 60 Db 178 CCAGCTAAATGATTTGCTCCCAATCTGCGCCAGATGGCGAAGAAGTTTCCCAAATGTACA 237 OY 61 PheLeuysVaLasPaVAlasPgLiLdeuasnservAlalaserLystrPGluilAsnAla 80      .....    :        :        Db 238 TTTCTTGAAGGTGGATGTGGATTTGAAGGCTTTGCTGTAGAGAGTGAATGTGAGGCA 297 OY 81 MetProthrPheValPheLeuLysLysGliLyslleilecglLysileValgIAlasp 100      .....    :        :



BASE COUNT 165 a 114 c 157 g 215 t 2 others  
ORIGIN

Alignment Scores:  
Pred. No.: 8.84e-48 Length: 653  
Score: 473.50 Matches: 90  
Percent Similarity: 88.33% Conservative: 16  
Best Local Similarity: 75.00% Mismatches: 10  
Query Match: 78.01% Indels: 4  
DB: 6 Gaps: 1

US-09-786-715-4 (1-118) x I38524 (1-653)

QY 3 SerSerGIUGLUGLIGLIVAlIleGLYCYSHisSerValAspGIUTrPlySGluInPhe 22  
|||||  
DB 69 TCATCCGAGGAGGACAAAGTGTCCGCTGCCCAAGAGTTGAGATGACAGACTTC 128  
23 GlnIysGlyValAspSerLysLysLeuValAlIleAspPheThrAlaSerTrpCysGly 42  
:::|  
DB 129 AAGAAAGCGGTTGACAACTAGAACTGGTGGTGCATTTACTGCTTCATGAGTGGCGS 188  
43 ProCysArgPheIleAlaProIleuAlaGluMetAlaLysThrProHisValIle 62  
|||||  
DB 189 CCTGCCGTTTATTATGCCCCAATCTCTGCTGACATGCTTAAGAAAGATGCCCATGTATTA 248  
QY 63 PheLeuLysValAspValAspGIUleuLysThrValAla-GluGIUphelLysValGIUAl 82  
|||||  
DB 249 TTCCTCAAGGTTGATGTGATGAACTGAAAGCTGTTACGCGGAAATGAGAGTGGAGGC 308  
QY 82 aMetProThrPheValPheLeuLysGIUGLIGLIVAlGluArgLeuValGIYAlaAr 102  
|||||  
DB 309 AATGCCAACTTTGCTTCATTAAGATGAAAGAAAGATGACAGAGTTGTTGGTGGCCA 368  
QY 102 gLySGluGIUleuGlnAlaThrValGluLysHisGIYAla-----IleThrAla 118  
:::|  
DB 369 GAAAGAGAGTTGACAGCAGACCATATGTAAGACATCTCTCTCTACTGTACTGCT 426

RESULT 9  
156999 I56999 Sequence 103 from patent US 5650505. 653 bp DNA Linear PAT 07-OCT-1997  
LOCUS DEFINITION I56999  
ACCESSION I56999  
VERSION I56999.1 GI:2477412  
KEYWORDS  
ORGANISM Unknown.  
SOURCE Unknown.  
REFERENCE 1 (bases 1 to 653)  
AUTHORS Ryals,J.A., Alexander,D.C., Beck,J.J., Duesing,J.H., Goodman,R.M.,  
Friedrich,L.B., Harms,C., Meins,F., Jr., Montoya,A. deceased,  
Moyer,M.B., Neuhaus,J.M., Payne,G.B., Spertsen,C., Stinson,J.R.,  
Ukres,S.J., Ward,E.R. and Williams,S.C.  
TITLE Chemically regulatable and anti-pathogenic DNA sequences and uses  
JOURNAL Patent: US 5650505-A 103 22-JUL-1997;  
FEATURES location/Qualifiers  
source 1..653  
BASE COUNT 165 a 114 c 157 g 215 t 2 others  
ORIGIN

Alignment Scores:  
Pred. No.: 8.84e-48 Length: 653  
Score: 473.50 Matches: 90  
Percent Similarity: 88.33% Conservative: 16  
Best Local Similarity: 75.00% Mismatches: 10  
Query Match: 78.01% Indels: 4  
DB: 6 Gaps: 1

US-09-786-715-4 (1-118) x I56999 (1-653)

QY 3 SerSerGIUGLUGLIGLIVAlIleGLYCYSHisSerValAspGIUTrPlySGluInPhe 22  
|||||  
DB 69 TCATCCGAGGAGGACAAAGTGTCCGCTGCCCAAGAGTTGAGATGACAGACTTC 128  
23 GlnIysGlyValAspSerLysLysLeuValAlIleAspPheThrAlaSerTrpCysGly 42  
:::|  
DB 129 AAGAAAGCGGTTGACAACTAGAACTGGTGGTGCATTTACTGCTTCATGAGTGGCGS 188  
43 ProCysArgPheIleAlaProIleuAlaGluMetAlaLysThrProHisValIle 62  
|||||  
DB 189 CCTGCCGTTTATTATGCCCCAATCTCTGCTGACATGCTTAAGAAAGATGCCCATGTATTA 248  
QY 63 PheLeuLysValAspValAspGIUleuLysThrValAla-GluGIUphelLysValGIUAl 82  
|||||  
DB 249 TTCCTCAAGGTTGATGTGATGAACTGAAAGCTGTTACGCGGAAATGAGAGTGGAGGC 308  
QY 82 aMetProThrPheValPheLeuLysGIUGLIGLIVAlGluArgLeuValGIYAlaAr 102  
|||||  
DB 309 AATGCCAACTTTGCTTCATTAAGATGAAAGAAAGATGACAGAGTTGTTGGTGGCCA 368  
QY 102 gLySGluGIUleuGlnAlaThrValGluLysHisGIYAla-----IleThrAla 118  
:::|

DB 69 TCATCCGAGGAGGACAAAGTGTCCGCTGCCCAAGAGTTGAGATGACAGACTTC 128  
QY 23 GlnIysGlyValAspSerLysLysLeuValAlIleAspPheThrAlaSerTrpCysGly 42  
:::|  
DB 129 AAGAAAGCGGTTGACAACTAGAACTGGTGGTGCATTTACTGCTTCATGAGTGGCGS 188  
QY 43 ProCysArgPheIleAlaProIleuAlaGluMetAlaLysThrProHisValIle 62  
|||||  
DB 189 CCTGCCGTTTATTATGCCCCAATCTCTGCTGACATGCTTAAGAAAGATGCCCATGTATTA 248  
QY 63 PheLeuLysValAspValAspGIUleuLysThrValAla-GluGIUphelLysValGIUAl 82  
|||||  
DB 249 TTCCTCAAGGTTGATGTGATGAACTGAAAGCTGTTACGCGGAAATGAGAGTGGAGGC 308  
QY 82 aMetProThrPheValPheLeuLysGIUGLIGLIVAlGluArgLeuValGIYAlaAr 102  
|||||  
DB 309 AATGCCAACTTTGCTTCATTAAGATGAAAGAAAGATGACAGAGTTGTTGGTGGCCA 368  
QY 102 gLySGluGIUleuGlnAlaThrValGluLysHisGIYAla-----IleThrAla 118  
:::|

RESULT 10  
159865 I59865 Sequence 103 from patent US 5654414. 653 bp DNA Linear PAT 07-OCT-1997  
LOCUS DEFINITION I59865  
ACCESSION I59865  
VERSION I59865.1 GI:2478497  
KEYWORDS  
ORGANISM Unknown.  
SOURCE Unknown.  
REFERENCE 1 (bases 1 to 653)  
AUTHORS Ryals,J.A., Beck,J.J. and Friedrich,L.B.  
TITLE Chemically inducible promoter of a cucumber chitinase/lysozyme gene  
JOURNAL Patent: US 5654414-A 103 05-AUG-1997;  
FEATURES location/Qualifiers  
source 1..653  
BASE COUNT 165 a 114 c 157 g 215 t 2 others  
ORIGIN

Alignment Scores:  
Pred. No.: 8.84e-48 Length: 653  
Score: 473.50 Matches: 90  
Percent Similarity: 88.33% Conservative: 16  
Best Local Similarity: 75.00% Mismatches: 10  
Query Match: 78.01% Indels: 4  
DB: 6 Gaps: 1

US-09-786-715-4 (1-118) x I59865 (1-653)

QY 3 SerSerGIUGLUGLIGLIVAlIleGLYCYSHisSerValAspGIUTrPlySGluInPhe 22  
|||||  
DB 69 TCATCCGAGGAGGACAAAGTGTCCGCTGCCCAAGAGTTGAGATGACAGACTTC 128  
QY 23 GlnIysGlyValAspSerLysLysLeuValAlIleAspPheThrAlaSerTrpCysGly 42  
:::|  
DB 129 AAGAAAGCGGTTGACAACTAGAACTGGTGGTGCATTTACTGCTTCATGAGTGGCGS 188  
43 ProCysArgPheIleAlaProIleuAlaGluMetAlaLysThrProHisValIle 62  
|||||  
DB 189 CCTGCCGTTTATTATGCCCCAATCTCTGCTGACATGCTTAAGAAAGATGCCCATGTATTA 248  
QY 63 PheLeuLysValAspValAspGIUleuLysThrValAla-GluGIUphelLysValGIUAl 82  
|||||  
DB 249 TTCCTCAAGGTTGATGTGATGAACTGAAAGCTGTTACGCGGAAATGAGAGTGGAGGC 308  
QY 82 aMetProThrPheValPheLeuLysGIUGLIGLIVAlGluArgLeuValGIYAlaAr 102  
|||||  
DB 309 AATGCCAACTTTGCTTCATTAAGATGAAAGAAAGATGACAGAGTTGTTGGTGGCCA 368  
QY 102 gLySGluGIUleuGlnAlaThrValGluLysHisGIYAla-----IleThrAla 118  
:::|

QY 43 ProCysArgPheIleAlaProIleLeuAlaGluMetAlaLysLysThrProHisValIle 62  
 DB 189 CTTGGCCGTTTATGTCACCAATTCCTTGGTGCATGCTAAGAAAGATGCCCATGTTATA 248  
 QY 63 PheLeuLysValAspValAspGluLysThrValAla-GluGluPheLysValGluAl 82  
 DB 249 TTCCTCAAGGTGATGTTGATGAAGACGTGTCAGCCGGGAATGGAGTGTGGAGGC 308  
 QY 82 ametProThrPheValPheLeuLysGluLysGluValGluArgLeuValGlyAlaAr 102  
 DB 309 AATGCCACTTTTGTCTTCAATTAAGATGGAAGAAGTGGACAGAGTGTGTGCCAA 368  
 QY 102 glysGluGluLeuGlnAlaThrValGluLysHisGlyAla-----IleThrAla 118  
 DB 369 GAAAGAGGAGTTGCACACACCATAGTGAAGCATGCTGCTCTGCTACTGTCACTGCT 426  
 RESULT 6  
 AR064647 653 bp DNA linear PAT 29-SEP-1999  
 LOCUS AR064647 653 bp DNA linear PAT 29-SEP-1999  
 DEFINITION Sequence 103 from patent US 5847258.  
 ACCESSION AR064647  
 VERSION AR064647.1 GI:5993955  
 KEYWORDS  
 SOURCE Unknown.  
 ORGANISM Unknown.  
 REFERENCE 1 (bases 1 to 653)  
 AUTHORS Ryals,J.A., Moyer,M.B., Payne,G.B. and Ward,E.R.  
 TITLE DNA encoding beta-1,3-glucanases  
 JOURNAL Patent: US 5847258-A 103 08-DEC-1998;  
 FEATURES  
 source 1. 653  
 BASE COUNT 165 a 114 c 157 g 215 t 2 others  
 ORIGIN  
 Alignment Scores:  
 Pred. No.: 8.84e-48 Length: 653  
 Score: 473.50 Matches: 90  
 Percent Similarity: 88.33% Conservative: 16  
 Best Local Similarity: 75.00% Mismatches: 10  
 Query Match: 78.01% Indels: 4  
 DB: 6 Gaps: 1  
 US-09-786-715-4 (1-118) x AR064647 (1-653)  
 QY 3 SerSerGluGluGlnValIleGlyCysHisSerValAspGluTrpLysGluGlnPhe 22  
 DB 69 TCATCCGAGAGGAGGACAAAGTGTTCGGCTGCCACAAAGTGAAGAAACGATCACTTC 128  
 QY 23 GlnLysGlyValAspSerLysLysLeuValIleAspPheThrAlaSerTrpCysGly 42  
 DB 129 AAGAAAGGCGTGGAGACTAAGAAACGTGGTGTGATTTTACTCTTCATGCTCGGS 188  
 QY 43 ProCysArgPheIleAlaProIleLeuAlaGluMetAlaLysLysThrProHisValIle 62  
 DB 189 CTTGGCCGTTTATGTCACCAATTCCTTGGTGCATGCTAAGAAAGATGCCCATGTTATA 248  
 QY 63 PheLeuLysValAspValAspGluLysThrValAla-GluGluPheLysValGluAl 82  
 DB 249 TTCCTCAAGGTGATGTTGATGAAGACGTGTCAGCCGGGAATGGAGTGTGGAGGC 308  
 QY 82 ametProThrPheValPheLeuLysGluLysGluValGluArgLeuValGlyAlaAr 102  
 DB 309 AATGCCACTTTTGTCTTCAATTAAGATGGAAGAAGTGGACAGAGTGTGTGCCAA 368  
 QY 102 glysGluGluLeuGlnAlaThrValGluLysHisGlyAla-----IleThrAla 118  
 DB 369 GAAAGAGGAGTTGCACACACCATAGTGAAGCATGCTGCTCTGCTACTGTCACTGCT 426  
 RESULT 7  
 AR067572

LOCUS AR067572 653 bp DNA linear PAT 29-SEP-1999  
 DEFINITION Sequence 103 from patent US 5851766.  
 ACCESSION AR067572  
 VERSION AR067572.1 GI:5998794  
 KEYWORDS  
 SOURCE Unknown.  
 ORGANISM Unknown.  
 REFERENCE 1 (bases 1 to 653)  
 AUTHORS Ryals,J.A. and Harms,C.  
 TITLE Process for isolating chemically regulatable DNA sequences  
 JOURNAL Patent: US 5851766-A 103 22-DEC-1998;  
 FEATURES  
 source 1. 653  
 BASE COUNT 165 a 114 c 157 g 215 t 2 others  
 ORIGIN  
 Alignment Scores:  
 Pred. No.: 8.84e-48 Length: 653  
 Score: 473.50 Matches: 90  
 Percent Similarity: 88.33% Conservative: 16  
 Best Local Similarity: 75.00% Mismatches: 10  
 Query Match: 78.01% Indels: 4  
 DB: 6 Gaps: 1  
 US-09-786-715-4 (1-118) x AR067572 (1-653)  
 QY 3 SerSerGluGluGlnValIleGlyCysHisSerValAspGluTrpLysGluGlnPhe 22  
 DB 69 TCATCCGAGAGGAGGACAAAGTGTTCGGCTGCCACAAAGTGAAGAAACGATCACTTC 128  
 QY 23 GlnLysGlyValAspSerLysLysLeuValIleAspPheThrAlaSerTrpCysGly 42  
 DB 129 AAGAAAGGCGTGGAGACTAAGAAACGTGGTGTGATTTTACTCTTCATGCTCGGS 188  
 QY 43 ProCysArgPheIleAlaProIleLeuAlaGluMetAlaLysLysThrProHisValIle 62  
 DB 189 CTTGGCCGTTTATGTCACCAATTCCTTGGTGCATGCTAAGAAAGATGCCCATGTTATA 248  
 QY 63 PheLeuLysValAspValAspGluLysThrValAla-GluGluPheLysValGluAl 82  
 DB 249 TTCCTCAAGGTGATGTTGATGAAGACGTGTCAGCCGGGAATGGAGTGTGGAGGC 308  
 QY 82 ametProThrPheValPheLeuLysGluLysGluValGluArgLeuValGlyAlaAr 102  
 DB 309 AATGCCACTTTTGTCTTCAATTAAGATGGAAGAAGTGGACAGAGTGTGTGCCAA 368  
 QY 102 glysGluGluLeuGlnAlaThrValGluLysHisGlyAla-----IleThrAla 118  
 DB 369 GAAAGAGGAGTTGCACACACCATAGTGAAGCATGCTGCTCTGCTACTGTCACTGCT 426  
 RESULT 8  
 I38524 653 bp DNA linear PAT 13-MAY-1997  
 LOCUS I38524 653 bp DNA linear PAT 13-MAY-1997  
 DEFINITION Sequence 103 from patent US 5614395.  
 ACCESSION I38524  
 VERSION I38524.1 GI:2084578  
 KEYWORDS  
 SOURCE Unknown.  
 ORGANISM Unknown.  
 REFERENCE 1 (bases 1 to 653)  
 AUTHORS Ryals,J.A., Alexander,D.C., Beck,J.J., Duesing,J.H., Goodman,R.M.,  
 Friedrich,L.B., Harms,C., Meins,F. Jr., Montoya,A. deceased,  
 Moyer,M.B., Neuhaus,J.-M., Payne,G.B., Sperisen,C., Stinson,J.R.,  
 Uknes,S.J., Ward,E.R. and Williams,S.C.  
 TITLE Chemically regulatable and anti-pathogenic DNA sequences and uses  
 JOURNAL Patent: US 5614395-A 103 25-MAR-1997;  
 FEATURES  
 source 1. 653  
 /organism="unknown"

AR020895	LOCUS	AR020895	653 bp	DNA	linear	PAT 05-DEC-1998
	DEFINITION	Sequence 103 from patent US 5789214.				
	ACCESSION	AR020895				
	VERSION	AR020895.1	GI:3975510			
	KEYWORDS					
	SOURCE	Unknown.				
	ORGANISM	Unknown.				
		Unclassified.				
	REFERENCE	1 (bases 1 to 653)				
	AUTHORS	Ryals, J.A., Friedrich, L.B., Uknes, S.J. and Ward, E.R.				
	TITLE	Method of inducing gene transcription in a plant				
	JOURNAL	Patent: US 5789214-A 103 04-AUG-1996;				
	FEATURES	Location/Qualifiers				
	Source	1..653				
		/organism="unknown"				
	BASE COUNT	165 a 114 c 157 g 215 t			2 others	
	ORIGIN					



Db	298	ATGCCAACTTTATTTCTCCTAAGAAGATGGAATAATTAGTACGACAAACACTGTGGTGGCAT	357
QY	101	LysValGlyLeu 104	
Db	358	AAAGATGGCTG 369	
RESULT 7			
LOCUS	B0888094		
DEFINITION	B0888094	612 bp	mrna linear EST 17-OCT-2002
	P003F02	populus petioles cDNA library	Populus tremula cDNA 5 prime,
ACCESSION	B0888094		RNA sequence.
VERSION	B0888094.1	GI:24099159	
KEYWORDS	EST.		
SOURCE	Populus tremula		
ORGANISM	Populus tremula		
	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;		
	Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids		
	; eurossids I; Malpighiales; Salicaceae; Populus.		
REFERENCE	1 (bases 1 to 612)		
AUTHORS	Umeberg, P., Bhalero, R., Jansson, S. and Sterky, F.		
TITLE	The poplar tree transcriptome: Analysis of expressed sequence tags		
	from multiple libraries		
JOURNAL	unpublished		
COMMENT	Contact: BHALERO RUPALI R.		
	Umea Plant Science Center		
	Department of Plant Physiology		
	University of Umea, 901 87 Umea, Sweden		
	Tel: +46 90 786 5279		
	Fax: +46 90 786 6676		
	Email: rupali.bhalero@plantphys.umu.se.		
FEATURES			
source	Location/Qualifiers		
	1..612		
	/organism="Populus tremula"		
	/mol_type="mRNA"		
	/db_xref="taxon:113636"		
	/issue_type="petioles"		
	/clone_lib="Populus petioles cDNA library"		
BASE COUNT	161 a 97 c 160 g 194 t		
ORIGIN			
Alignment Scores:			
Pred. No.:	3,59e-39	Length:	612
Score:	385.00	Matches:	70
Percent Similarity:	78.85%	Conservative:	12
Best Local Similarity:	67.31%	Mismatches:	22
Query Match:	60.44%	Indels:	0
DB:	13	Gaps:	0
US-09-786-715-2 (1-122) x B0888094 (1-612)			
QY	1	MetAlaGluGluGlyGlnValIleAlaCysHisLysIleAspGluTrpGluGlyGlnLeu	20
Db	56	ATGGCGCAAGAGCAGCATGATTGGCCGCACACAGCGATACCTGGAAAGACATTTTC	115
QY	21	GlyLysTrpLysAspSerGluLysLeuValValAlaAspPheThrAlaSerTrpCysGly	40
Db	116	GAGAAAGGAGAAAGGCTCTCAGAAACTGATTTCTCGTGGATTTTACTGCTTCATGGGTCCA	175
QY	41	ProCysArgAlaIleAlaProTyrPheThrGluLeuAlaLysAsnAsnProAsnValAla	60
Db	176	CCAGTGAATAAGATGTCCTCCCAATCTGCCGACAGTTGGCGAAGAGAGTTCCCTAATGTCCA	235
QY	61	PheLeuLysValAspValAspGluLeuLysAsnSerValAlaSerLysTrpGluIleAsnAla	80
Db	236	TTCTTGAAGGCGATGTCGATGATTTGAAGCGCTGTCTGTCGAGAGTGGAATGTGAGGCA	295
QY	81	MetProThrPheValPheLeuLysGlyLysIleIleGluLysIleValAlaLysAsp	100
Db	296	ATGCCAACTTTATTTCTCCTAAGAAGATGGAATAATTAGTACGACAAACACTGTGGTGGCAT	355
QY	101	LysValGlyLeu 104	

DB	356	AAAGTGGCCG 367
RESULT 8		
B0830685		
LOCUS	B0830685	615 bp mRNA linear EST 15-OCT-2002
DEFINITION	T011F03 Populus apical shoot cDNA library	Populus tremula x Populus tremuloides cDNA 5 prime, mRNA sequence.
ACCESSION	B0830685	
VERSION	B0830685.1	GI:24008183
KEYWORDS	EST.	
SOURCE	Populus tremula x Populus tremuloides	
ORGANISM	Populus tremula x Populus tremuloides	
REFERENCE	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids; eurosids I; Malpighiales; Salicaceae; Populus.	
AUTHORS	1 (bases 1 to 615)	
TITLE	Umeberg, P., Bhalerao, R.R., Jansson, S. and Sterky, F.	
JOURNAL	The poplar tree transcriptome: Analysis of expressed sequence tags from multiple libraries	
COMMENT	Unpublished	
CONTACT	BHALERAO RUPALI R.	
UNMA PLANT SCIENCE CENTER		
DEPARTMENT OF PLANT PHYSIOLOGY		
UNIVERSITY OF UMEA, 901 87 UMEA, SWEDEN		
TEL: +46 90 786 5279		
FAX: +46 90 786 6676		
EMAIL: rupali.bhalerao@plantphys.umu.se.		
FEATURES	Location/Qualifiers	
SOURCE	1..615	
	/organism="Populus tremula x Populus tremuloides"	
	/mol_type="mRNA"	
	/db_xref="taxon:47664"	
	/tissue_type="apical shoot"	
	/clone_lib="Populus apical shoot cDNA library"	
BASE COUNT	154 a 97 c 161 g 203 t	
ORIGIN		
Alignment Scores:		
Pred. No.:	3,61e-39	Length: 615
Score:	385.00	Matches: 70
Percent Similarity:	78.85%	Conservative: 12
Best Local Similarity:	67.31%	Mismatches: 22
Query Match:	60.44%	Indels: 0
DB:	13	Gaps: 0
US-09-786-715-2 (1-122) x B0830685 (1-615)		
QY	1	Meklaaglugluglgnvalilalealcyshtsllysilaaspgultrpgluglgnleu 20
DB	35	ATGCCCGAAGGACACAGATTATGCTGCCACACAGATGATCTGGAAGACATTTC 94
QY	21	glylstrpplrysaspserglulysleuvalvalaspherhralasertpocslg 40
DB	95	GAGAAAGGAAAGGGGCTCAGAACTGATTCGTGCGATTTTACTGCTTCATGCGTCA 154
QY	41	Procyasagalaiallealaprtyrphethrlnleualalysasnasrproasvalala 60
DB	155	CCATGTAAATGATTCGCCAATCTTCGCCAGTTGGGAAAGATTCCATAATGTCCA 214
QY	61	PhleulysvalasvalaspgluleuanserValalaserlSTRPGLulcasnala 80
DB	215	TTCCTGAAGGAGGATGTGATGATTTGAAGCTCTGTCTGAGAGTGGAATGTGGAGCA 274
QY	81	MetProthrPhevalPheleulysclgyllysleilegIulysllevalGlyalasp 100
DB	275	ATGCCAATTTTATTTCTCTGAAGATGGAAATTAATGACAAACATGTGGTGTGAT 334
QY	101	LysValGlyleu 104
DB	335	AAAGTGTCTG 346
RESULT 9		
B0873533		

LOCUS B0873533 641 bp mRNA linear EST 16-OCT-2002  
DEFINITION 0056f01 Populus flower cDNA library Populus balsamifera subsp.  
trichocarpa cDNA 5 prime, mRNA sequence.  
ACCESSION B0873533  
VERSION B0873533.1 GI:24065057  
KEYWORDS EST.  
SOURCE Populus balsamifera subsp. trichocarpa  
ORGANISM Populus balsamifera subsp. trichocarpa  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids  
; eucosids I; Malpighiales; Salicaceae; Populus.  
REFERENCE 1 (bases 1 to 641)  
AUTHORS Umeberg, P., Bhalerao, R.R., Jansson, S. and Sterky, F.  
TITLE The poplar tree transcriptome: Analysis of expressed sequence tags  
from multiple libraries  
JOURNAL Unpublished  
COMMENT Contact: BHALERAO RUPALI R.  
Umea Plant Science Center  
Department of Plant Physiology  
University of Umea, 901 87 Umea, Sweden  
Tel: +46 90 786 5279  
Fax: +46 90 786 6676  
Email: rupali.bhalerao@plantphys.umu.se.  
location/Qualifiers  
1. 641  
/organism="Populus balsamifera subsp. trichocarpa"  
/mol\_type="mRNA"  
/db\_xref="taxon:3694"  
/clone\_lib="Populus flower cDNA library"  
/note="Organ: flower"  
BASE COUNT 163 a 102 c 175 g 201 t  
ORIGIN  
Alignment Scores:  
Pred. No.: 3,82e-39 Length: 641  
Score: 385.00 Matches: 70  
Percent Similarity: 78.85% Conservative: 12  
Best Local Similarity: 67.31% Mismatches: 22  
Query Match: 60.44% Indels: 0  
Gaps: 0  
US-09-786-715-2 (1-122) x B0873533 (1-641)  
QY 1 MetAlaGlUGlUGlInValIleAlaCySHsLysIleAspGluTrpGluGlnLeu 20  
|||||  
DB 86 ATGGCGGAGAAAGCAAGTATTGCGTCCACACAGTGATACCTGGAAAGCAATTTC 145  
QY 21 GlyLysTrpLysAspSerGluLysLeuValValAlaSpPheThrAlaSerTrpCysGly 40  
||| |||  
DB 146 GAGAAAGGAAAGCGGCTCGAAGAACTGATGTCGTGATTTTACTGCTTCATGTCGCA 205  
QY 41 ProCysArgAlaIleAlaProTyrPheThrGluLeuAlaLysAsnSerProAsnValAla 60  
|||||  
DB 206 CCATGTAAATGATCTCTCCATCTTCCGCGAGTGGCGAAGATTTCCTAATGTCACA 265  
QY 61 PheLeuLysValAspValAspGluLeuAsnSerValAlaSerLysTrpGluIleAsnAla 80  
|||||  
DB 266 TTCTTGAAGGTGATGTGATGAATGAAGCGCTGTCGAGAGATGGAATGCGAGCA 325  
QY 81 MetProThrPheValPheLeuLysGlyLysIleIleGluLysIleValGlyAlaAsp 100  
|||||  
DB 326 ATGCCCAACTTATTCTCTGAAAGATGGAATAATTAGTGACAAACTGTGGTGCTGAT 385  
QY 101 LysValGlyLeu 104  
||| |||||  
DB 386 AAAAGATGGCCTC 397  
RESULT 10 520 bp mRNA linear EST 17-OCT-2002  
LOCUS B0885783  
DEFINITION R036A12 Populus root cDNA library Populus tremula x Populus  
tremuloides cDNA 5 prime, mRNA sequence.

ACCESSION B0885783  
VERSION B0885783.1 GI:24077300  
KEYWORDS EST.  
SOURCE Populus tremula x Populus tremuloides  
ORGANISM Populus tremula x Populus tremuloides  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids  
; eucosids I; Malpighiales; Salicaceae; Populus.  
REFERENCE 1 (bases 1 to 520)  
AUTHORS Umeberg, P., Bhalerao, R.R., Jansson, S. and Sterky, F.  
TITLE The poplar tree transcriptome: Analysis of expressed sequence tags  
from multiple libraries  
JOURNAL Unpublished  
COMMENT Contact: BHALERAO RUPALI R.  
Umea Plant Science Center  
Department of Plant Physiology  
University of Umea, 901 87 Umea, Sweden  
Tel: +46 90 786 5279  
Fax: +46 90 786 6676  
Email: rupali.bhalerao@plantphys.umu.se.  
location/Qualifiers  
1. 520  
/organism="Populus tremula x Populus tremuloides"  
/mol\_type="mRNA"  
/db\_xref="taxon:47664"  
/tissue\_type="root"  
/clone\_lib="Populus root cDNA library"  
BASE COUNT 131 a 89 c 138 g 162 t  
ORIGIN  
Alignment Scores:  
Pred. No.: 5.21e-39 Length: 520  
Score: 383.00 Matches: 70  
Percent Similarity: 80.77% Conservative: 14  
Best Local Similarity: 67.31% Mismatches: 20  
Query Match: 60.13% Indels: 0  
Gaps: 0  
US-09-786-715-2 (1-122) x B0885783 (1-520)  
QY 1 MetAlaGlUGlUGlInValIleAlaCySHsLysIleAspGluTrpGluGlnLeu 20  
|||||  
DB 27 ATGGCAGAAAGAAAGCAAGTATGCTCCACACCGTGTGATCTCGGAAAGCAATTTC 86  
QY 21 GlyLysTrpLysAspSerGluLysLeuValValAlaSpPheThrAlaSerTrpCysGly 40  
||| |||  
DB 87 GAGAAAGGAAAGGCACTCGAAGCTGATGTGTCGTGATTTTACTGCTTCATGTCCT 146  
QY 41 ProCysArgAlaIleAlaProTyrPheThrGluLeuAlaLysAsnSerProAsnValAla 60  
|||||  
DB 147 CCATGTAAATGATCTCTCCGCAATTTTCGCGGATTTGGCCAAAGATTCCACATGTCACC 206  
QY 61 PheLeuLysValAspValAspGluLeuAsnSerValAlaSerLysTrpGluIleAsnAla 80  
|||||  
DB 207 TTCTTGAAGGTGATGTGATGAATGAAGCGCTGTCGCGAGTGGAGATGCGAGCG 266  
QY 81 MetProThrPheValPheLeuLysGlyLysIleIleGluLysIleValGlyAlaAsp 100  
|||||  
DB 267 ATGCCCAACTTATTCTCTGAAAGATGGAATAATTAGTGACAAATTTGTGGTGCTGAT 326  
QY 101 LysValGlyLeu 104  
||| |||||  
DB 327 AAAAGATGGCCTC 338  
RESULT 11 523 bp mRNA linear EST 15-OCT-2002  
LOCUS B0812532  
DEFINITION UL97TE03 Populus leaf cDNA library Populus tremula x Populus  
tremuloides cDNA 5 prime, mRNA sequence.  
ACCESSION B0812532  
VERSION B0812532.1 GI:23967453  
KEYWORDS EST.  
SOURCE Populus tremula x Populus tremuloides

REFERENCE	1 (bases 1 to 523)
UNNEBERG, P., BHALERAO, R.R., JANSSON, S. and STERKY, F.	
TITLE	the poplar tree transcriptome: Analysis of expressed sequence tags from multiple libraries
JOURNAL	unpublished
COMMENT	Contact: BHALERAO RUPALI R. Umea Plant Science Center Department of Plant Physiology University of Umea, 901 87 Umea, Sweden Tel: +46 90 786 5279 Fax: +46 90 786 6676 Email: rupali.bhalerao@plantphys.umu.se.
FEATURES	Location/Qualifiers
SOURCE	1..523
	/organism="Populus tremula x Populus tremuloides"
	/mol_type="mRNA"
	/db_xref="taxon:47664"
	/tissue_type="apical shoot"
	/clone_lib="Populus apical shoot cDNA library"
BASE COUNT	138 a 92 c 138 g 155 t
ORIGIN	
Alignment Scores:	
Pred. No.:	5.25e-39
Score:	383.00
Percent Similarity:	80.77%
Best Local Similarity:	67.31%
Query Match:	60.13%
DB:	13 Gaps: 0
US-09-786-715-2 (1-122) x BU829684 (1-523)	
QY	1 MetaaaglugluglglglnvallealeaAcyshtslyslleaspjgltptpglglglnleu 20.
DB	59 ATGCACAGAGAGAGACCAAGTATTCGCCGCACACCGCTGATGTCGGAAAGACCAATTC 118
QY	21 GlytsttptylsaspsersglnlyslleuvalvalyaspnethrAlAsertTpcysgly 40
DB	119 GAGAGAGGAGAAAGGACCTGACAGACCTGATTTGGTGCGATTTTACTGCTTCATGGTGCTT 178
QY	41 ProCysatgaAlaileAlaProTyrPhehrGlnleuAlaAlaYsAsnAsProAsnValAla 60
DB	179 CCATGTAATTCATTCGCGCCCAATTTCCGGCATTTGGCCAGAGATTCACCAATGTCACC 238
QY	61 PheleuysValaspsvalaspglnleuanserValAlAserLysTTPgluITleasnaIa 80
DB	239 TTCTTGAGAGGTGACGCTGATGATTAAGCCTGTTGCTGCGGAGTGGCAAGTGCAGGCG 298
QY	81 MetProthrPhevalPheleuLysLysLysIlelelgulysIlevalGlyAlaasp 100
DB	299 ATGCCCACTTTATTTTCCGTAAGAAGCGGAAATTAAGGACAAAAATTGCGTGCCTGAT 358
QY	101 LysValGlyLeu 104
DB	359 AAAGATGGCCTC 370
RESULT 13	
LOCUS	B1136756 531 bp mRNA linear EST 31-DEC-2001
DEFINITION	F073P30Y Populus flower cDNA library Populus balsamifera subsp. trichocarpa cDNA, mRNA sequence.
ACCESSION	B1136756
VERSION	B1136756.1 GI:18017704
KEYWORDS	EST.
SOURCE	Populus balsamifera subsp. trichocarpa
ORGANISM	Populus balsamifera subsp. trichocarpa
REFERENCE	Unkari, P., Tegel, T., Gustafsson, P., Bhalerao, R., Jansson, S., Nilsson, O., Sundberg, B., Nilsson, P., Uhlen, M., Sandberg, G. and

TITLE  
JOURNAL  
COMMENT  
Lundberg, J.  
Gene expression in Populus  
Unpublished  
Contact: Eriandsson R  
Department of Biotechnology  
Royal Institute of Technology  
Teknikringen 30, Stockholm S-10044, Sweden  
Tel: 46 8 790 8287  
Fax: 46 8 245452  
Email: riker1@biochem.kth.se.

FEATURES  
source  
1.531  
/organism="Populus balsamifera subsp. trichocarpa"  
/mol\_type="mRNA"  
/sub\_species="trichocarpa"  
/db\_xref="taxon:3694"  
/clone\_lib="Populus flower cDNA library"  
/note="organ: flower"

BASE COUNT  
ORIGIN  
140 a 84 c 147 g 160 t

Alignment Scores:  
Pred. No.: 5.36e-39 Length: 531  
Score: 383.00 Matches: 70  
Percent Similarity: 80.77% Conservative: 14  
Best Local Similarity: 67.31% Mismatches: 20  
Query Match: 60.13% Indels: 0  
DB: 12 Gaps: 0

US-09-786-715-2 (1-122) x B1136756 (1-531)

QY 1 MetAlaGlUGlUGlGlnValIleAlaCYSHsLYsIleAspGluTrpGluGlnLeu 20  
|||||  
52 ATGGCAGAGAAAGAGCAAGTATGCGTCGCCACACCGATTCGTGGAAGAGCAATTC 111  
|||||

QY 21 GtLYsTrpLYsAspSerGluLYsLeuValValAlaSPheThrAlaSerTrpCYsGly 40  
|||||  
112 GAGAAAGGAAAGAGGAGCTCAGAGCTGATGCTGCTGATTTACTCTCATGCTCCCT 171  
|||||

QY 41 ProCYsArgAlaIleAlaProTYrPheThrGluLeuAlaLYsAsnSnpProAsnValAla 60  
|||||  
172 CCATGTAATTCATTCGCGCAGTTTCGGGATTTGCCAAGAAATTCACCATGTCACC 231  
|||||

QY 61 PheLeuLYsValAspValAspGluLeuAsnSerValAlaSerLYsTrpGluIleAsnAla 80  
|||||  
232 TTCTTCAAGGTGACCTGATGTAATGAAGCCCTGCTCGCGAGTGGAGAGCGCG 291  
|||||

QY 81 MetProThrPheValPheLeuLYsLYsIleIleGluLYsIleValGlyAlaasp 100  
|||||  
292 ATGCCAAGCTTTTATTTCTCGAAGATGGAAATTAAGTGACAAATTTGGGTCTGAT 351  
|||||

QY 101 LysValGlyLeu 104  
|||  
352 AAGATGGCCTC 363  
|||

RESULT 14  
BU884977 532 bp mRNA linear EST 17-OCT-2002  
LOCUS BU884977  
DEFINITION R018H02 Populus root cDNA library Populus tremula x Populus  
tremuloides cDNA 5 prime, mRNA sequence.  
ACCESSION BU884977  
VERSION BU884977.1 GI:24076494  
KEYWORDS  
SOURCE  
ORGANISM  
Populus tremula x Populus tremuloides  
Populus tremula x Populus tremuloides  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids  
; eurosids I; Malpighiales; Salicaceae; Populus.

REFERENCE  
AUTHORS  
TITLE  
JOURNAL  
Unneberg, P., Bhalerao, R.R., Jansson, S. and Sterky, F.  
The poplar tree transcriptome: Analysis of expressed sequence tags  
from multiple libraries  
Unpublished

COMMENT  
Contact: BHALERAO RUPALI R.  
Umea Plant Science Center  
Department of Plant Physiology  
University of Umea, 901 87 Umea, Sweden  
Tel: +46 90 786 5279  
Fax: +46 90 786 6676  
Email: rupali.bhalerao@plantphys.umu.se.

FEATURES  
source  
1.532  
/organism="Populus tremula x Populus tremuloides"  
/mol\_type="mRNA"  
/db\_xref="taxon:47664"  
/tissue\_type="root"  
/clone\_lib="Populus root cDNA library"

BASE COUNT  
ORIGIN  
139 a 90 c 141 g 162 t

Alignment Scores:  
Pred. No.: 5.37e-39 Length: 532  
Score: 383.00 Matches: 70  
Percent Similarity: 80.77% Conservative: 14  
Best Local Similarity: 67.31% Mismatches: 20  
Query Match: 60.13% Indels: 0  
DB: 13 Gaps: 0

US-09-786-715-2 (1-122) x BU884977 (1-532)

QY 1 MetAlaGlUGlUGlGlnValIleAlaCYSHsLYsIleAspGluTrpGluGlnLeu 20  
|||||  
43 ATGGCAGAGAAAGAGCAAGTATGCGTCGCCACACCGTTCGTGGAAGAGCAATTC 102  
|||||

QY 21 GtLYsTrpLYsAspSerGluLYsLeuValValAlaSPheThrAlaSerTrpCYsGly 40  
|||||  
103 GAGAAAGGAAAGAGGAGCTCAGAGCTGATGCTGCTGATTTACTCTCATGCTCCCT 162  
|||||

QY 41 ProCYsArgAlaIleAlaProTYrPheThrGluLeuAlaLYsAsnSnpProAsnValAla 60  
|||||  
163 CCATGTAATTCATTCGCGCAGTTTCGGGATTTGCCAAGAAATTCACCAATGTCACC 222  
|||||

QY 61 PheLeuLYsValAspValAspGluLeuAsnSerValAlaSerLYsTrpGluIleAsnAla 80  
|||||  
223 TTCTTCAAGGTGACCTGATGTAATGAAGCCCTGCTCGCGAGTGGAGAGCGCG 282  
|||||

QY 81 MetProThrPheValPheLeuLYsLYsIleIleGluLYsIleValGlyAlaasp 100  
|||||  
283 ATGCCAAGCTTTTATTTCTCGAAGAGGAAATTAAGTGACAAATTTGGGTCTGAT 342  
|||||

QY 101 LysValGlyLeu 104  
|||  
343 AAGATGGCCTC 354  
|||

RESULT 15  
BU827985 544 bp mRNA linear EST 15-OCT-2002  
LOCUS BU827985  
DEFINITION K014P03P Populus apical shoot cDNA library Populus tremula x  
Populus tremuloides cDNA 5 prime, mRNA sequence.  
ACCESSION BU827985  
VERSION BU827985.1 GI:24002337  
KEYWORDS  
SOURCE  
ORGANISM  
Populus tremula x Populus tremuloides  
Populus tremula x Populus tremuloides  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids  
; eurosids I; Malpighiales; Salicaceae; Populus.

REFERENCE  
AUTHORS  
TITLE  
JOURNAL  
Unneberg, P., Bhalerao, R.R., Jansson, S. and Sterky, F.  
The poplar tree transcriptome: Analysis of expressed sequence tags  
from multiple libraries  
Unpublished  
Contact: BHALERAO RUPALI R.  
Umea Plant Science Center  
Department of Plant Physiology  
University of Umea, 901 87 Umea, Sweden





GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - nucleic search, using frame\_plus\_p2n model

Run on: August 17, 2003, 17:46:53 ; Search time 165.884 Seconds

(without alignments)  
1920.219 Million cell updates/sec

Title: US-09-786-715-4

Perfect score: 607

Sequence: 1 MASSEGEQVIGCHSDVEMKE.....VGARKEELQATVEKHGATTA 118

Scoring table:

BLOSUM62	
Xgapop 10.0 , Ygapext 0.5	
Ygapop 10.0 , Xgapext 0.5	
Fgapop 6.0 , Fgapext 7.0	
Delop 6.0 , Delext 7.0	

Searched: 2552756 segs, 1349719017 residues

Total number of hits satisfying chosen parameters: 5105512

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Command line parameters:

-MODEL=frame+ p2n.model -DEV=rlp  
-O=cgna\_1/USPTO.spool/US09786715/runat\_11082003\_150513\_6025/app\_query.fasta.1.1052  
-DB=N\_Geneseq.19jun03 -QMT=fastap -SUFFIX=ring -MINMATCH=0.1 -LOOPCL=0  
-LOOPEXT=0 -UNITS=bits -START=1 -END=-1 -MATRIX=blomsum2 -TRANS=human40.cdl  
-LIST=45 -LOCALIGN=200 -THR.SCORE=pct -THR.MAX=100 -THR.MIN=0 -ALIGN=15  
-MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000  
-USER=US09786715.GCGN.1.1.874 -tunat\_11082003\_150513\_6025 -NCPU=6 -ICPU=3  
-NO\_MMAP -LARGESUBSTRY -NEG\_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG  
-DEV.TIMEOUT=120 -WARN.TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6  
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

N\_Geneseq.19jun03:\*

- 1: /SIDSL/gcgdata/geneseq/geneseqn\_emb1/NA1980.DAT:\*
- 2: /SIDSL/gcgdata/geneseq/geneseqn\_emb1/NA1981.DAT:\*
- 3: /SIDSL/gcgdata/geneseq/geneseqn\_emb1/NA1982.DAT:\*
- 4: /SIDSL/gcgdata/geneseq/geneseqn\_emb1/NA1983.DAT:\*
- 5: /SIDSL/gcgdata/geneseq/geneseqn\_emb1/NA1984.DAT:\*
- 6: /SIDSL/gcgdata/geneseq/geneseqn\_emb1/NA1985.DAT:\*
- 7: /SIDSL/gcgdata/geneseq/geneseqn\_emb1/NA1986.DAT:\*
- 8: /SIDSL/gcgdata/geneseq/geneseqn\_emb1/NA1987.DAT:\*
- 9: /SIDSL/gcgdata/geneseq/geneseqn\_emb1/NA1988.DAT:\*
- 10: /SIDSL/gcgdata/geneseq/geneseqn\_emb1/NA1989.DAT:\*
- 11: /SIDSL/gcgdata/geneseq/geneseqn\_emb1/NA1990.DAT:\*
- 12: /SIDSL/gcgdata/geneseq/geneseqn\_emb1/NA1991.DAT:\*
- 13: /SIDSL/gcgdata/geneseq/geneseqn\_emb1/NA1992.DAT:\*
- 14: /SIDSL/gcgdata/geneseq/geneseqn\_emb1/NA1993.DAT:\*
- 15: /SIDSL/gcgdata/geneseq/geneseqn\_emb1/NA1994.DAT:\*
- 16: /SIDSL/gcgdata/geneseq/geneseqn\_emb1/NA1995.DAT:\*
- 17: /SIDSL/gcgdata/geneseq/geneseqn\_emb1/NA1996.DAT:\*
- 18: /SIDSL/gcgdata/geneseq/geneseqn\_emb1/NA1997.DAT:\*
- 19: /SIDSL/gcgdata/geneseq/geneseqn\_emb1/NA1998.DAT:\*
- 20: /SIDSL/gcgdata/geneseq/geneseqn\_emb1/NA1999.DAT:\*
- 21: /SIDSL/gcgdata/geneseq/geneseqn\_emb1/NA2000.DAT:\*
- 22: /SIDSL/gcgdata/geneseq/geneseqn\_emb1/NA2001.DAT:\*
- 23: /SIDSL/gcgdata/geneseq/geneseqn\_emb1/NA2001A.DAT:\*
- 24: /SIDSL/gcgdata/geneseq/geneseqn\_emb1/NA2002.DAT:\*
- 25: /SIDSL/gcgdata/geneseq/geneseqn\_emb1/NA2003.DAT:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed,

and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	607	100.0	574	21 AA251738	Catalpa speciosa t
2	473.5	78.0	653	16 AA099783	plant SAR gene pl.
3	473.5	78.0	653	20 AA062799	Tobacco SAR CHX in
4	473.5	78.0	653	20 AAV81683	Tobacco protein-sy
5	446	73.5	601	21 AA251740	Glycine max thioire
6	437	72.0	392	25 ABX18045	Human GDP-mannose
7	434	71.5	509	22 ABX21664	Human GDP-mannose
8	433.5	71.4	502	22 AAH87768	Pepermint plant O.
9	430	70.8	390	25 ABX19403	Human GDP-mannose
10	430	70.8	738	21 AA251739	Glycine max thioire
11	421	69.4	419	21 AAA31942	Plant microsatelli
12	420	69.2	3888	24 ABSS3097	DNA encoding thior
13	420	69.2	3888	24 ABN89581	Phascolin promoter
14	419	69.0	345	24 AB212358	Arabidopsis thalia
15	419	69.0	345	24 ABN89587	Arabidopsis thalia
16	419	69.0	470	21 AAC37781	Arabidopsis thalia
17	419	69.0	561	21 AAC51522	Arabidopsis thalia
18	419	69.0	563	21 AAC34121	Arabidopsis thalia
19	419	69.0	3129	24 ABSS3095	DNA encoding thior
20	419	69.0	3129	24 ABN89579	Phaseolin promoter
21	419	69.0	3888	24 ABSS3096	DNA encoding Oleos
22	419	69.0	3888	24 ABN89580	Phaseolin promoter
23	414	68.2	4935	24 ABN89586	Promoter-Oleosin t
24	401	66.1	614	21 AA251741	Vernonia mespilifo
25	392	64.6	560	21 AAC41961	Arabidopsis thalia
26	391	64.4	524	21 AAC33829	Arabidopsis thalia
27	390	64.3	686	15 AAQ78205	Gene coding for pr
28	390	64.3	687	21 AAC66375	Rice thioiredoxin h
29	389	64.1	357	24 AB213931	Arabidopsis thalia
30	389	64.1	480	21 AAC36542	Arabidopsis thalia
31	389	64.1	563	21 AAC34211	Arabidopsis thalia
32	388	64.1	652	21 AAC48656	Arabidopsis thalia
33	388	63.9	353	21 AAA31118	Plant microsatelli
34	386	63.6	572	21 AAA31222	Plant microsatelli
35	382	62.9	572	21 AAC52069	Arabidopsis thalia
36	377	62.1	576	25 ABX56868	Arabidopsis thalia
37	374.5	61.7	360	24 AB212359	Arabidopsis thalia
38	374	61.6	328	21 AAA31097	Plant microsatelli
39	373	61.4	285	21 AAA31225	Plant microsatelli
40	369	60.8	647	21 AAC38520	Arabidopsis thalia
41	368.5	60.7	590	21 AAC38792	Arabidopsis thalia
42	368	60.6	346	21 AAA31096	Plant microsatelli
43	367	60.5	297	21 AAA31940	Plant microsatelli
44	364	60.0	320	21 AAA31785	Plant microsatelli
45	362	59.6	341	21 AAA31924	Plant microsatelli

## ALIGNMENTS

RESULT 1	AA251738
ID	AA251738 standard; CDNA; 574 BP.
XX	AA251738;
XX	
DT	04-JUL-2000 (first entry)
XX	
DE	Catalpa speciosa thioiredoxin CDNA.
XX	
KW	Catalpa speciosa thioiredoxin; clone ncs.pk0010.e3; chimeric gene;
KW	transgenic plant; seed storage protein; allergenicity; ss.
XX	
OS	Catalpa speciosa.
XX	
Key	Location/Qualifiers
FT	CDS 63..419

```
FT FT      /*tag= a
XX XX      /product= "thioredoxin"
PN PN      MO200014239-A2.
XX XX      16-MAR-2000.
XX XX      07-SEP-1999; 99WO-US20420.
XX XX      08-SEP-1998; 98US-0099501.
XX XX      (DUPO ) DU PONT DE NEMOURS & CO E. I.
XX XX      Allen SM, Thorpe CJ, Lu AL.
XX XX      WPI; 2000-256987/22.
DR DR      P-PsDB; AAY70480.
XX XX
PT PT      New isolated polynucleotide encoding thioredoxin polypeptide is useful
PT PT      for producing transgenic plants with an altered level of thioredoxin -
XX XX
XX XX      Claim 3; Page 29; 33pp; English.
XX XX
XX XX      The present cDNA sequence encodes Catalpa speciosa thioredoxin
XX XX      protein. The cDNA was derived from clone ncs-PK010.e3, which was
XX XX      isolated from a cDNA library prepared from C. speciosa developing seed
XX XX      tissue. Chimeric genes encoding all or a portion of the thioredoxin
XX XX      protein, in sense or antisense orientation are constructed, wherein
XX XX      expression of the chimeric gene results in production of altered levels
XX XX      of the thioredoxin protein in a transformed host cell. Thioredoxin is
XX XX      involved in the disassembly of seed storage proteins during germination
XX XX      by reducing S-S bonds and in the bread making process. Over expression of
XX XX      thioredoxin in cereals may reduce the allergenicity of cereal proteins
XX XX      protein engineered into cereal crops with high sulfhydryl content.
XX XX
SQ SQ      Sequence 574 BP; 171 A; 91 C; 157 G; 155 T; 0 other;

Alignment Scores:
Pred. No.: 6,3e-71 Length: 574
Score: 607.00 Matches: 118
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 21 Gaps: 0

US-09-786-715-4 (1-118) x AA251738 (1-574)
OY 1 MetAlaSerSerGluGluGluGluValIleGlyCysHisSerValAspGluTrpGlyGlu 20
DB 63 ATGGCTTCTTCAGAAAGGACAAAGATCGGTGGCTCCCTCGACGAGTGGAAAGGAG 122
OY 21 GlnPheGlnLysGlyValAspSerLysLysLeuValValIleAspPheThrAlaSerTrp 40
DB 123 CAGTTCACAAAGAGGTTCGACTGACTAGAAACTGCTGTAATAGACTTCACGGCTTCTCTG 182
OY 41 CysGlyProCysArgPheIleAlaProIleLeuAlaGluMetAlaLysLysThrProHis 60
DB 183 TCGCGACATCGCGCTTTCATTCGCTCAATCTTGCGTCGATGGCCAAAGACACACCCAT 242
OY 61 ValIlePheLeuLysValAspValAspGluLeuLysThrValAlaGluGluPheLysVal 80
DB 243 GTCATATTCCTGAATGACGACGATGATCAACACATGCTGCTGAGGAATTCCAAGTG 302
OY 81 GlnAlaMetProThrPheValPheLeuLysGluGlyLysGluValGluArgLeuValGly 100
DB 303 GAGGCTATCCCAACCTTCGTGTTCCTCAAGGAAGGAAAGAAAGCTTGTGGGA 362
OY 101 AlaArgLysGluGluLeuGlnAlaThrValGluLysHisGlyAlaIleThrAla 118
DB 363 GCAAGGAAGGAGGAATTCGAGGCCACAGAGTTGAGAAACATGGGCTATCACTGCT 416
RESULT 2
AA099783
```

```
ID ID      AA099783 standard; DNA; 653 BP.
XX XX
XX XX      AA099783;
AC AC
XX XX      24-JUN-1996 (first entry)
DT DT
XX XX      Plant SAR gene pl.4.3.
XX XX
XX XX      SAR; tobacco; protein-synthesis independent gene; cyclohexamide;
XX XX      systemic acquired resistance response; pl.4.3; anti-pathogen;
XX XX      plant protection; ss.
XX XX
XX XX      Nicotiana acuminata.
XX XX
XX XX      WO9519443-A2.
XX XX
XX XX      20-JUL-1995.
XX XX
XX XX      03-JAN-1995; 95WO-IB00002.
XX XX
XX XX      13-JAN-1994; 94US-0181271.
XX XX
XX XX      (CIBA ) CIBA GEIGY AG.
XX XX
XX XX      Alexander DC, Ryals JA, Uknes SJ, Ward ER;
XX XX
XX XX      WPI; 1995-263872/34.
XX XX
XX XX      New DNA contg. plant systemic acquired resistance genes - and
XX XX      transgenic plants contg. them, impart disease and pest resistance,
XX XX      also Atradiopsis gene promoter to control DNA transcription
XX XX
XX XX      Disclosure; Page 61; 85pp; English.
XX XX
XX XX      This sequence represents the DNA sequence of a tobacco protein-synthesis
XX XX      independent gene. The gene is involved in the regulation of the
XX XX      systemic acquired resistance (SAR) response. This gene is designated
XX XX      pl.4.3. This sequence and AA099784 represent SAR genes that are not used
XX XX      in the recombinant/chimaeric DNA molecules of the invention.
XX XX      AA099784-099790 and AA099806 are SAR genes used in the
XX XX      recombinant/chimaeric DNA molecules of the invention. The wild type
XX XX      genes corresponding to these sequences can all be chemically induced in a
XX XX      plant in a protein-synthesis independent manner. SAR genes are involved
XX XX      in the protection of plants against pests and disease. These sequences
XX XX      were isolated by differential screening of a cDNA library, followed by
XX XX      analysis by Northern hybridisation to RNA in the presence and absence of
XX XX      cyclohexamide. The genes are used in the creation of transgenic plants.
XX XX      Transgenic expression of 2 or more of the recombinant molecules of the
XX XX      invention that encode anti-pathogenic proteins provides a synergistic
XX XX      increase in plant protection, and may also offer protection against a
XX XX      wider range of pathogens.
XX XX
SQ SQ      Sequence 653 BP; 165 A; 114 C; 157 G; 215 T; 2 other;

Alignment Scores:
Pred. No.: 3,73e-53 Length: 653
Score: 473.50 Matches: 90
Percent Similarity: 88.33% Conservative: 16
Best Local Similarity: 75.00% Mismatches: 10
Query Match: 78.01% Indels: 4
DB: 16 Gaps: 1

US-09-786-715-4 (1-118) x AAQ99783 (1-653)
OY 3 SerSerGluGluGlyGlnValIleGlyCysHisSerValAspGluTrpGlyGluGlnPhe 22
DB 69 TCATCCGAGGAGGAGCAAGTGTTCGCTCCACCAAGTTGAGGAAGCAAGTACTTC 128
OY 23 GlnLysGlyValAspSerLysLysLeuValValIleAspPheThrAlaSerTrpCysGly 42
DB 129 AAGAAAGGCGCTTGAGACTAAGAAACTGCTGCTGCTGCTTACTTCTTCATGCTCGGS 188
OY 43 ProCysArgPheIleAlaProIleLeuAlaGluMetAlaLysLysThrProHisValIle 62
```



```

Db      |||||
189 CCTGCCGTTTATGCCCCCAATCTGCTGACATTGCTAAGACATGCCCATGTTATA 248
QY      63 PheLeuLySValAspValAspGluLeuLySThrValAlaA-GluGluPheLySValGluAl 82
Db      249 TTCTCCAAAGGTTGATGTTGATGACTGTAAGACTGTTTACGCGGAAATGAGTGTGAGAGC 308
QY      82 aMePProThrPheValPheLeuLySGluGlyLySGluValGluATGLeuValGlyAlaAr 102
Db      309 AATGCCAATTGTTGCTTCTCATTAAGATGAAAGAGAGTGCACAGATGTTGTGGCCAA 368
QY      102 gLySGluGluLeuGlnAlaThrValGluLyShiSGlyAlaA-----lIeThraAl 118
Db      369 GAAAGAGGAGTTGCAGCAGACCATATGTAAGACATGCTGCTCTGCTACTGTCATGCT 426

RESULT 3
AAV62799 standard; cDNA; 653 BP.
XX
AC      AAV62799;
XX
DT      05-MAR-1999 (first entry)
XX
DE      Tobacco SAR CHX independent gene clone 1.4.3.
XX
KW      Chemically regulatable DNA promoter; expression control; pesticide;
KM      herbicide tolerance; systemic acquired resistance gene; ss.
XX
OS      Nicotiana acuminata.
XX
PN      US5851766-A.
XX
PD      22-DEC-1998.
XX
PF      31-MAY-1995; 95US-0456262.
XX
PR      31-MAY-1995; 95US-0456262.
XX
PA      (NOVS ) NOVARTIS FINANCE CORP.
XX
PI      Harms C, Ryals JA;
XX
DR      WPI; 1999-08036/07.
XX
PT      Isolating chemically regulatable DNA sequences in plants - useful
XX      for chemically controlling expression in transformed plants
XX
PS      Example 40A; Column 257-260; 175pp; English.
XX
CC      This sequence represents a clone of the tobacco systemic acquired
CC      resistance (SAR) cyclohexamide (CHX) independent.
CC      This gene can be isolated using the method of the invention.
CC      The method is for isolating a chemically regulatable DNA promoter
CC      fragment from the 5' flanking region of a chemically regulatable gene in
CC      a plant tissue. The method allows isolation of sequences which will be
CC      useful for the controlled expression of genes, under the control of a
CC      non-coding regulatable sequence. This is useful in plants with a
CC      herbicide or pesticide detoxification mechanism under the control of a
CC      chemical regulator, the regulator being applied before or with the
CC      herbicide or pesticide to give optimal tolerance. The promoter fragment
CC      is useful for controlling sequences which encode traits such as
CC      height, shape, development, male or female sterility, and the ability
CC      of the plant to withstand cold, heat, salt and drought. The chemical
CC      induction of the promoter allows the regulation of production of
CC      compounds, e.g. flavours, fragrances, pigments, natural sweeteners,
CC      industrial feedstocks, antimicrobials and pharmaceuticals, by
CC      biosynthesis or metabolite conversion, whose biosynthesis is controlled
CC      by endogenous or foreign genes. The method allows control over the time
CC      and rate of gene expression either throughout the whole plant, or in
CC      localized tissues, to achieve e.g. fungal or insect resistance by for
CC      instance dusting the leaves with the chemical regulator. Controlling the
CC      developmental processes by the application of a regulating chemical in
CC      e.g. the commercial production of cultivated crops allows processes such

```

```

CC      as germination, flower formation and fruit ripening to be synchronised at
CC      a given time.
XX
SQ      Sequence 653 BP; 165 A; 114 C; 157 G; 215 T; 2 other;
XX
Alignment Scores:
Pred. No.: 3,73e-53 Length: 653
Score: 473.50 Matches: 90
Percent Similarity: 88.33% Conservative: 16
Best Local Similarity: 75.00% Mismatches: 10
Query Match: 78.01% Indels: 4
DB: Gaps: 1

US-09-786-715-4 (1-118) x AAV62799 (1-653)
QY      3 SerSergLugLugLyGlnValIlleGlyCySHiServalAspGluTrpLySGluInphe 22
Db      69 TCATCCGAGAGGACGACAAAGTGTTCGCTCCCAAGAGTGTGAGCAATGCAACGACTACTTC 128
QY      23 GlnLySGlyValAspSerLySLeuValValIlleAspPheThrAlaSerTrpCysGly 42
Db      129 AAGAAAGGCGTTGACACTAAGAACTGGTGTGCTGATTTTACTGCTTCATGTCGCGGS 188
QY      43 ProCyASrPheIlleAlaProIlleLeuAlaGluMetAlaLySLeuThrProHisValIlle 62
Db      189 CCTGCCGTTTATGCCCCCAATCTGCTGACATTGCTCAACAGATGCCCATGTTATA 248
QY      63 PheLeuLySValAspValAspGluLeuLySThrValAlaA-GluGluPheLySValGluAl 82
Db      249 TTCTCCAAAGGTTGATGTTGATGACTGTAAGACTGTTTACGCGGAAATGAGTGTGAGAGC 308
QY      82 aMePProThrPheValPheLeuLySGluGlyLySGluValGluATGLeuValGlyAlaAr 102
Db      309 AATGCCAATTGTTGCTTCTCATTAAGATGAAAGAGAGTGCACAGATGTTGTGGCCAA 368
QY      102 gLySGluGluLeuGlnAlaThrValGluLyShiSGlyAlaA-----lIeThraAl 118
Db      369 GAAAGAGGAGTTGCAGCAGACCATATGTAAGACATGCTGCTCTGCTACTGTCATGCT 426

RESULT 4
AAV81683 standard; DNA; 653 BP.
XX
AC      AAV81683;
XX
DT      25-FEB-1999 (first entry)
XX
DE      Tobacco protein-synthesis independent gene pl.4.3.
XX
KW      Regulation; transcription; plant tissue; chimeric construction; PR;
KM      pathogenesis-related protein; anti-pathogenic; transgenic plant;
KW      beta-1,3-glucanase activity; pest resistance; ss.
XX
OS      Nicotiana sp.
XX
PN      US5847258-A.
XX
PD      08-DEC-1998.
XX
PF      31-MAY-1995; 95US-0457364.
XX
PR      31-MAY-1995; 95US-0457364.
XX
PR      08-MAR-1998; 89US-0165667.
XX
PR      06-FEB-1988; 89US-0305566.
XX
PR      24-MAR-1989; 89US-0329018.
XX
PR      20-JUN-1989; 89US-0326672.
XX
PR      07-SEP-1989; 89US-0425504.
XX
PR      07-SEP-1990; 90US-0580431.
XX
PR      21-DEC-1990; 90US-0632441.
XX
PR      01-APR-1991; 91US-0678378.
XX
PR      27-SEP-1991; 91US-0768122.
XX
PR      06-MAR-1992; 92US-0848506.
XX
PR      06-NOV-1992; 92US-0973197.

```

PR 06-APR-1993; 93US-0042847.  
PR 12-APR-1993; 93US-0045957.  
PR 16-JUL-1993; 93US-0093301.  
PR 13-JAN-1994; 94US-0181271.  
XX  
PA (NOVS ) NOVARTIS FINANCE CORP.  
XX  
PI Moyer MB, Payne GB, Ryals JA, Ward ER;  
XX WPI: 1999-059180/05.  
XX  
PT DNA encoding pathogenesis-related glucanase proteins - useful for  
PT producing transgenic plants with enhanced disease or pest resistance  
XX  
PS Example 40; Column 253-254; 169pp: English.  
XX  
CC The present invention describes a DNA molecule encoding a  
CC pathogenesis-related (PR) protein having beta-1,3-glucanase activity  
CC selected from PR-2, PR-2', PR-2'', PR-N, PR-O and PR-O'. Also described  
CC are: (i) a chimeric gene comprising the above DNA molecule linked to a  
CC heterologous promoter; (ii) a vector containing the chimeric gene;  
CC (iii) a host cell containing the chimeric gene; (iv) a transgenic plant  
CC containing the chimeric gene; and (v) a seed from the transgenic plant.  
CC The DNA molecule is used to produce transgenic plants with enhanced  
CC disease or pest resistance. The present sequence represents a tobacco  
CC protein-synthesis independent gene p1.4.3 from the present invention.  
XX  
SQ Sequence 653 BP; 165 A; 114 C; 157 G; 215 T; 2 other:  
  
Alignment Scores:  
Pred. No.: 3.73e-53 Length: 653  
Score: 473.50 Matches: 90  
Percent Similarity: 88.33% Conservative: 16  
Best Local Similarity: 75.00% Mismatches: 10  
Query Match: 78.01% Indels: 4  
DB: 20 gaps: 1  
  
US-09-786-715-4 (1-118) x AAV81683 (1-653)  
QY 3 SerSerGluGluGluGluValIleGlyCysHisSerValAspGluTrpLysGluGluPhe 22  
Db 69 TCATCCGAGGAGGAGACAGCTTCGGCTGCCACAGTTGAGCATGAGACAGACTCTTC 128  
QY 23 GlnLysGlyValAspSerLysLysLeuValValIleAspPheThrAlaSerTrpCysGly 42  
Db 129 AAGAAAGCCCTTGAAGTAAAGAAAGCTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 188  
QY 43 ProCysArpPheIleAlaProIleLeuAlaGluMetAlaLysLysTrpProHisValIle 62  
Db 189 CCTTCCCGTTTATGCCCCAATTCCTGCTGACATTGCTAAGAAAGATGCCCATGTTATA 248  
QY 63 PheLeuLysValAspValAspGluLeuLysThrValAla-GluGluPheLysValGluAl 82  
Db 249 TTCCTCAGGTTATGTTGATGAGACGAGACGCTTTCACGGGAGATGGAGTGTGAGGC 308  
QY 82 aMetProThrPheValPheLeuLysGluGluValGluAlaGluArgLeuValGluAla 102  
Db 309 AAGCCAACTTTTGTCTTCAATTAAAGATGAGAAAGAGTGCACAGATGTTGTTGCCAA 368  
QY 102 GlySGluGluLeuGluAlaThrValGluLysHisGlyAla-----IleThrAla 118  
Db 369 GAAAGAGGAGTTGACACAGACACATGATGAGCATGCTGCTCTCTACTGTCTACTGCT 426  
  
RESULT 5  
AA251740  
ID AA251740 standard; cDNA: 601 BP.  
XX  
AC AA251740;  
XX  
DT 04-JUL-2000 (first entry)  
XX  
DE Glycine max thioredoxin cDNA-2.  
XX

KW Glycine max thioredoxin; clone sfil.pk0029.e2; chimeric gene; soybean;  
KW transgenic plant; seed storage protein; allergenicity; ss.  
XX  
OS Glycine max.  
XX  
FH Key Location/Qualifiers  
FT CDS 37..408  
FT /\*tag= a  
FT /product= "thioredoxin"  
XX  
PN WO200014239-A2.  
XX  
PD 16-MAR-2000.  
XX  
PF 07-SEP-1999; 99WO-US20420.  
XX  
PR 08-SEP-1998; 98US-0099501.  
XX  
PA (DUPO ) DU PONT DE NEMOURS & CO E I.  
XX  
PI Allen SM, Thorpe CJ, Lu AL;  
XX  
DR WPI: 2000-256987/22.  
DR P-PSDB: AAY70482.  
XX  
PT New isolated polynucleotide encoding thioredoxin polypeptide is useful  
PT for producing transgenic plants with an altered level of thioredoxin -  
XX  
PS Claim 3; Page 30-31; 33pp: English.  
XX  
CC The present cDNA sequence encodes glycine max thioredoxin protein. The  
CC cDNA was derived from clone sfil.pk0029.e2, which was isolated from a  
CC cDNA library prepared from soybean immature flower. Chimeric genes  
CC encoding all or a portion of the thioredoxin protein, in sense or  
CC antisense orientation are constructed, wherein expression of the chimeric  
CC gene results in production of altered levels of the thioredoxin protein  
CC in a transformed host cell. Thioredoxin is involved in the disassembly of  
CC seed storage proteins during germination by reducing S-S bonds and in the  
CC bread making process. Over expression of thioredoxin in cereals may  
CC reduce the allergenicity of any transgenic protein engineered into  
CC cereal crops with high sulfhydryl content.  
XX  
SQ Sequence 601 BP; 194 A; 103 C; 136 G; 168 T; 0 other:  
  
Alignment Scores:  
Pred. No.: 1.47e-49 Length: 601  
Score: 446.00 Matches: 80  
Percent Similarity: 85.59% Conservative: 21  
Best Local Similarity: 67.80% Mismatches: 17  
Query Match: 73.48% Indels: 0  
DB: 21 gaps: 0  
  
US-09-786-715-4 (1-118) x AA251740 (1-601)  
QY 1 MetAlaSerSerGluGluGluValIleGlyCysHisSerValAspGluTrpLysGlu 20  
Db 37 ATGGCTGAAGTGAAGAGGAGGAGCTACGCGCTCCACACCGTTGATGAGTGAAGCTG 96  
QY 21 GlnPheGlnLysLysValAspSerLysLysLeuValValIleAspPheThrAlaSerTrp 40  
Db 97 CAACCTCCAGAAAGCAAAAGAGCTCAAAAATGATTTGGTGGATTTTACTGCTCTCG 156  
QY 41 CysGlyProCysArpPheIleAlaProIleLeuAlaGluMetAlaLysLysTrpProHis 60  
Db 157 TGTGTCATTCGCGTTTATGCGCCAGTTCTTCGAGATGTGCAGAAAGAACTCCTGA 216  
QY 61 ValIlePheLeuLysValAspValAspGluLeuLysThrValAlaGluLysVal 80  
Db 217 TTGATCTTCCTCAAGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 276  
QY 81 GluAlaMetProThrPheValPheLeuLysGluGluValGluAlaGluArgLeuValGly 100  
Db 277 GAGGCCATCCCAACCTTCTCTTGTGAAGATGGCGGAGATGCTGGACAAAGGTGGTGGT 336

OY 101 AlaArgLySGluLeuGlnAlaThrValGluLySHSGlyAlaIleThrAla 118  
||| |||:||||| |||: ||| |||  
Db 337 GCTAGTAAGATGACCTTCACGCCACCATAGCCAGCATGCTCTTGGT 390

RESULT 6  
ABX18045  
ID ABX18045 standard; cDNA; 392 BP.  
XX  
AC ABX18045;  
XX  
DT 10-FEB-2003 (first entry)  
XX  
DE Human GDP-mannose 4,6-dehydratase (GM4,6D) DNA #102.  
XX  
KW Human: GDP-mannose 4,6-dehydratase; GM4,6D; gene; ss; inflammation;  
KW cellular fucosylation; glycoconjugate fucosylation; transplant rejection;  
KW arthritis; asthma; sepsis; reperfusion injury; stroke; infection;  
KW complex carbohydrate; gene replacement therapy; immunosuppressive;  
KW antiinflammatory; antiarthritic; antibacterial; cerebroprotective;  
KW antiasthmatic; vasotropic.  
XX  
OS Homo sapiens.  
XX  
PN US2002110548-A1.  
XX  
PD 15-AUG-2002.  
XX  
PF 11-JUN-2001; 2001US-0878574.  
XX  
PR 22-NOV-1996; 96US-0753233.  
PR 03-DEC-1997; 97US-0984246.  
PR 09-SEP-1998; 98US-0149674.  
PR 14-JUN-1999; 99US-0333177.  
XX  
PA (GENM ) GENETICS INST INC.  
XX  
PI Sullivan F, Kriz R, Kumar R;  
XX  
DR WPI: 2003-066673/06.  
XX  
PT New composition comprising GDP-mannose 4,6-dehydratase (GM4,6D)  
PT peptide, for manufacturing complex carbohydrates, or as targets for  
PT screening GM4,6D antagonists for treating e.g. arthritis, or transplant  
PT rejection  
XX  
PS Disclosure; SEQ ID NO 104; 6pp; English.  
XX  
CC The invention relates to a composition comprising a human GDP-mannose  
CC 4,6-dehydratase (GM4,6D) peptide. The peptide is useful for identifying  
CC GM4,6D inhibitors. GM4,6D inhibitors are useful for reducing inflammation  
CC in a mammalian subject and for treating or ameliorating diseases affected  
CC by the level of cellular fucosylation or diseases affected by the  
CC fucosylation of glycoconjugates. These diseases include arthritis,  
CC transplant rejection, asthma, sepsis, reperfusion injury, stroke or  
CC infection. The GM4,6D peptide or a polynucleotide encoding it is also  
CC useful for manufacturing complex carbohydrates and as targets for  
CC screening small molecule antagonists of the activity of the enzyme. The  
CC polynucleotide is useful in developing an assay for defects in the  
CC enzyme, as well as in gene replacement therapy. Sequences  
CC ABX17942-ABX17944 and ABX17947-ABX33716 represent DNA molecules encoding  
CC human GM4,6D peptides of the invention.  
CC Note: The sequence data for this patent did not form part of the printed  
CC specification but was obtained in electronic format directly from USPTO  
CC at [seqdata.uspto.gov/sequence.html](http://seqdata.uspto.gov/sequence.html).  
XX  
SQ Sequence 392 BP; 99 A; 80 C; 107 G; 106 T; 0 other;

Alignment Scores:  
Pred. No.: 1,28e-48 Length: 392  
Score: 437.00 Matches: 77  
Percent Similarity: 86.09% Conservative: 22  
Best Local Similarity: 66.96% Mismatches: 16

Query Match: 71.99% Indels: 0  
DB: 25 Gaps: 0  
US-09-786-715-4 (1-118) x ABX18045 (1-392)

OY 4 SerGluGluGlyGlnValIleGlyCysSHSerValAspGluTrpLySGluInpHeGln 23  
:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||  
Db 9 GCTGAAGAGGCTCAGGTTCTCGCGCCGCCACCGTTGATGCGTGGACACGACGACGAG 68  
OY 24 LysGlyValAspSerLysLysLeuValValIleAspPheThrAlaSerTrpCysGlyPro 43  
||| |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||  
Db 69 AATGGAAAGACTCCCAAAAATTGATTGGTGATTTTACTGCTCTCGTGCGTGCATCA 128  
OY 44 CysArgPheIleAlaProIleLeuValIleGluMetAlaLysThrProHisValIlePhe 63  
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||  
Db 129 TGCCGTTTATTTGCCCCAGCTTCTGGCGAGATTGCTAGCGATCTCTCAAGATCTTC 188  
OY 64 LeuLysValAspValAspGluLeuLysThrValAlaGluLysPheLysValGluAlaMet 83  
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||  
Db 189 CTCACAGGTGATGTGATGATGAGAGCGCTGTTGCTGAGGAAATATTCATGAGGCGATG 248  
OY 84 ProThrPheValPheLeuLysGluGlyLysGluValGluArgLeuValAlaArgLys 103  
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||  
Db 249 CCACCTTCTCTCTTCTTGAAGATGCAAGATCGTGAATAGGTTGTTGCTCTAGAGAG 308  
OY 104 GluGluLeuGlnAlaThrValGluLySHSGlyAlaIleThrAla 118  
||| |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||  
Db 309 GAGGAGCTGCACACTCACCATGACCAAGCATGATCTGCTGCTGCT 353

RESULT 7  
ABX21664  
ID ABX21664 standard; cDNA; 402 BP.  
XX  
AC ABX21664;  
XX  
DT 10-FEB-2003 (first entry)  
XX  
DE Human GDP-mannose 4,6-dehydratase (GM4,6D) DNA #3721.  
XX  
XX  
KW Human: GDP-mannose 4,6-dehydratase; GM4,6D; gene; ss; inflammation;  
KW cellular fucosylation; glycoconjugate fucosylation; transplant rejection;  
KW arthritis; asthma; sepsis; reperfusion injury; stroke; infection;  
KW complex carbohydrate; gene replacement therapy; immunosuppressive;  
KW antiinflammatory; antiarthritic; antibacterial; cerebroprotective;  
KW antiasthmatic; vasotropic.  
XX  
OS Homo sapiens.  
XX  
PN US2002110548-A1.  
XX  
PD 15-AUG-2002.  
XX  
PF 11-JUN-2001; 2001US-0878574.  
XX  
PR 22-NOV-1996; 96US-0753233.  
PR 03-DEC-1997; 97US-0984246.  
PR 09-SEP-1998; 98US-0149674.  
PR 14-JUN-1999; 99US-0333177.  
XX  
PA (GENM ) GENETICS INST INC.  
XX  
PI Sullivan F, Kriz R, Kumar R;  
XX  
DR WPI: 2003-066673/06.  
XX  
PT New composition comprising GDP-mannose 4,6-dehydratase (GM4,6D)  
PT peptide, for manufacturing complex carbohydrates, or as targets for  
PT screening GM4,6D antagonists for treating e.g. arthritis, or transplant  
PT rejection  
XX  
PS Disclosure; SEQ ID NO 3723; 6pp; English.  
XX  
XX The invention relates to a composition comprising a human GDP-mannose



XX DE Human GDP-mannose 4,6-dehydratase (GM4,6D) DNA #1460.  
XX XX  
XX KW Human; GDP-mannose 4,6-dehydratase; GM4,6D; gene; ss; inflammation;  
XX KW cellular fucosylation; glycoconjugate fucosylation; transplant rejection;  
XX KW arthritis; asthma; sepsis; reperfusion injury; stroke; infection;  
XX KW complex carbohydrate; gene replacement therapy; immunosuppressive;  
XX KW antiinflammatory; antiarthritic; antibacterial; cerebroprotective;  
XX KW antihistaminic; vasotropic.  
XX OS Homo sapiens.  
XX PN US2002110548-A1.  
XX PD 15-AUG-2002.  
XX XX  
XX PF 11-JUN-2001; 2001US-0878574.  
XX XX  
XX PR 22-NOV-1996; 96US-0753233.  
XX PR 03-DEC-1997; 97US-0984246.  
XX PR 09-SEP-1998; 98US-0149674.  
XX PR 14-JUN-1999; 99US-0333177.  
XX PA (GENY ) GENETICS INST INC.  
XX PI Sullivan F, Kriz R, Kumar R;  
XX DR WPI, 2003-066673/06.  
XX XX  
XX PT New composition comprising GDP-mannose 4,6-dehydratase (GM4,6D)  
XX PT peptide, for manufacturing complex carbohydrates, or as targets for  
XX PT screening GM4,6D antagonists for treating e.g. arthritis, or transplant  
XX PT rejection  
XX PS Disclosure: SEQ ID NO 1462; fpp; English.  
XX XX  
XX CC The invention relates to a composition comprising a human GDP-mannose  
XX CC 4,6-dehydratase (GM4,6D) peptide. The peptide is useful for identifying  
XX CC GM4,6D inhibitors. GM4,6D inhibitors are useful for reducing inflammation  
XX CC in a mammalian subject and for treating or ameliorating diseases affected  
XX CC by the level of cellular fucosylation or diseases affected by the  
XX CC fucosylation of glycoconjugates. These diseases include arthritis,  
XX CC transplant rejection, asthma, sepsis, reperfusion injury, stroke or  
XX CC infection. The GM4,6D peptide or a polynucleotide encoding it is also  
XX CC useful for manufacturing complex carbohydrates and as targets for  
XX CC screening small molecule antagonists of the activity of the enzyme. The  
XX CC polynucleotide is useful in developing an assay for defects in the  
XX CC enzyme, as well as in gene replacement therapy. Sequences  
XX CC ABX17942-ABX17944 and ABX17947-ABX3716 represent DNA molecules encoding  
XX CC human GM4,6D peptides of the invention.  
XX CC Note: The sequence data for this patent did not form part of the printed  
XX CC specification but was obtained in electronic format directly from USPTO  
XX CC at seqdata.uspto.gov/sequence.html.  
XX SQ Sequence 390 BP; 97 A; 86 C; 97 G; 110 T; 0 other;  
  
Alignment Scores:  
Pred. No.: 1.08e-47 Length: 390  
Score: 430.00 Matches: 77  
Percent Similarity: 88.18% Conservative: 20  
Best Local Similarity: 70.00% Mismatches: 13  
Query Match: 70.84% Indels: 0  
DB: 25 Gaps: 0  
  
US-09-786-715-4 (1-118) x ABX19403 (1-390)  
OY 1 MetAlaSerSerGluGluGlyValIleGlyCysHisSerValAlaSpGluTrpLysGlu 20  
DB 57 ATGGCTGAAGTGGAGAGGACAGGTCAATCGGCGCCACACCGTTGATGAGTGGAGCTG 116  
OY 21 GlnPheGlnLysGlyValAlaSpSerLysLysLeuValIleAspPheThrAlaSerTrp 40  
DB 117 CAATCTCAGAAATGCAAAAGACATCCAAAAAAGCTGATTGGTGGTGGTATTACTGCTTCCTGG 176

OY 41 CysGlyProCysArgPheIleAlaProIleLeuAlaGluMetAlaLysTrpProHis 60  
DB 177 TGTGATCCATGCGCGTTTATGGCCCGAGTTCTTGAGAGATTCGCAAGAAACCTCTGAA 236  
OY 61 ValIlePheLeuLysValAlaSpValAlaSpGluLeuLysTrpValAlaGluLupheLysVal 80  
DB 237 TTGATCTTCCTCAAAAGTGGATGTGATGAGAGAGCCGTGCTGAGAGAAATATTCATT 296  
OY 81 GluAlaMetProThrPheValAlaPheLeuLysGluGlyValGluValGluArgLeuValGly 100  
DB 297 GAGGCCATGCGCAACCTTCCTCTTGAAGATGCGAGATGCGACAGAGGTGGTGGT 356  
OY 101 AlaArgLysGluGluLeuGluAlaThrVal 110  
DB 357 GCTAGTAAAGATGACCTTCACGCCACATTA 386  
  
RESULT 10  
AAZ51739  
ID AAZ51739 standard; cDNA; 738 BP.  
XX  
XX AAZ51739;  
AC  
XX  
XX 04-JUL-2000 (first entry)  
XX DE Glycine max thioedoxin cDNA-1.  
XX XX  
XX KW Glycine max thioedoxin; clone sahic.pk001.117; chimeric gene; soybean;  
XX KW transgenic plant; seed storage protein; allergenicity; ss.  
XX XX  
XX Glycine max.  
XX  
XX  
XX Key Location/Qualifiers  
XX CDS 68..430  
XX FT /\*tag= a  
XX FT /product= "Thioedoxin"  
XX FT  
XX PN WO200014239-A2.  
XX PD 16-MAR-2000.  
XX PF 07-SEP-1999; 99WO-US20420.  
XX PR 08-SEP-1998; 98US-0099501.  
XX PA (DUPO ) DU PONT DE NEMOURS & CO E I.  
XX PI Allen SM, Thorpe CJ, Lu AL;  
XX DR WPI: 2000-256987/22.  
XX DR P-PSDB: AAT70481.  
XX PT New isolated polynucleotide encoding thioedoxin polypeptide is useful  
XX PT for producing transgenic plants with an altered level of thioedoxin -  
XX PS Claim 3; Page 30; 33pp; English.  
XX XX  
XX CC The present cDNA sequence encodes Glycine max thioedoxin protein. The  
XX CC cDNA was derived from clone sahic.pk001.117, which was isolated from a  
XX CC cDNA library prepared from soybean tissue sprayed with authority  
XX CC herbicide. Chimeric genes encoding all or a portion of the thioedoxin  
XX CC protein, in sense or antisense orientation are constructed, wherein  
XX CC expression of the chimeric gene results in production of altered levels  
XX CC of the thioedoxin protein in a transformed host cell. Thioedoxin is  
XX CC involved in the disassembly of seed storage proteins during germination  
XX CC by reducing S-S bonds and in the bread making process. Over expression of  
XX CC thioedoxin in cereals may reduce the allergenicity of any transgenic  
XX CC protein engineered into cereal crops with high sulfhydryl content.  
XX SQ Sequence 738 BP; 260 A; 120 C; 146 G; 212 T; 0 other;  
  
Alignment Scores:  
Pred. No.: 2.57e-47 Length: 738

Score: 430.00 Matches: 77  
Percent Similarity: 87.61% Conservative: 22  
Best Local Similarity: 68.14% Mismatches: 14  
Query Match: 70.84% Indels: 0  
DB: 21 Gaps: 0

US-09-786-715-4 (1-118) x AA251739 (1-738)

OY 3 SerSerGlucIuGlucIuValIleGlyCysHisSerValAspGluTrpLysGluGlnPhe 22  
|||||  
DB 77 TCATCGAAGAGGACAGCAAGTCATTAGCTGCCACACCTTGAGAGATGAGCAATCAATC 136  
OY 23 GlnLysGlyValAspSerLysLysLeuValValIleAspPheThrAlaSerTrpCysGly 42  
|||||  
DB 137 CAGAGGGCAGCAACATTCAGAAACCTCATGTTGTGATTTTACTGCTTGGTGGA 196  
OY 43 ProCysArpPheIleAlaProIleLeuAlaGluMetLysLysThrProHisValIle 62  
|||||  
DB 197 CCAAGCCGTTTCATGACCATTTCTGTGCTGAGCTGAGCTAAGAGATTCCACAAGTGTGATA 256  
OY 63 PheLeuLysValAspValAspGluLeuLysThrValAlaGluGluPheLysValGluAla 82  
|||||  
DB 257 TTCTTAAGGTGATGTGAGCAATTAAGAGTGTTCACAGATTGGGCTATTGAGGCT 316  
OY 83 MetProThrPheValPheLeuLysGluGlyLysGluValGluArgLeuValGlyAlaArg 102  
|||||  
DB 317 ATGCCCACTTTTGTGTTGTGAAGAGGAGCAAGCTTGTGAGCAAGTGTGGAGCAAG 376  
OY 103 LysGluGluLeuGlnAlaThrValGluLysHisGlyAla 115  
|||||  
DB 377 AAGGATGAGCTGCAGCAGAAATACAGAAACATGTGCT 415

RESULT 11  
AAA31942  
ID AAA31942 standard; DNA: 419 BP.

XX  
AC AAA31942:  
DT 05-JUL-2000 (first entry)  
XX  
DE Plant microsatellite marker #903.  
XX  
KW Plant microsatellite sequence; core repeat sequence; detection; probe;  
KW DNA polymorphism; genome mapping; physical mapping; fingerprinting;  
KW variety identification; genetic variability evaluation; primer; ss.  
XX  
OS Eucalyptus grandis.  
XX  
PN W09967421-A1.  
XX  
PD 29-DEC-1999.  
XX  
PF 25-JUN-1999; 99WO-NZ00092.  
XX  
PR 25-JUN-1998; 98US-0105307.  
XX  
PA (GENE-) GENESIS RES & DEV CORP LTD & FLETCHER.  
PA (FLET-) FLETCHER CHALLENGE FORESTS LTD.  
XX  
PI Havukkala IF, Bloksberg LN, Glenn M;  
XX  
DR WPI; 2000-116958/10.  
XX  
XX  
PT New plant microsatellite markers and associated flanking species for  
PT the detection of polymorphic genetic markers -  
XX  
XX  
PS Claim 1; Page 342; 392pp; English.

XX  
XX Sequences AAA31040-A32093 represent novel plant microsatellite sequences  
XX and associated flanking species. The sequences comprise a central core  
XX repeat sequence, especially selected from the sequences AAA32094-A32096  
XX with left and right flanking sequences. The polynucleotide sequences  
XX can be used in the detection of DNA polymorphisms, in genome mapping,

CC in physical mapping, in positional cloning of genes, in variety  
CC identification and in evaluation of genetic variability within and  
CC between plant tissues, populations, cultivars, species and species  
CC groups. They may also be used to design hybridization probes for  
CC oligonucleotide fingerprinting and library screening and to design  
CC primers for microsatellite-primed PCR. Microsatellite markers are  
CC useful to locate specific economically useful genes in plant genomes.  
XX

SO Sequence 419 BP; 108 A; 83 C; 129 G; 96 T; 3 other;

Alignment Scores:  
Pred. No.: 1.86e-46 Length: 419  
Score: 421.00 Matches: 81  
Percent Similarity: 86.36% Conservative: 14  
Best Local Similarity: 73.64% Mismatches: 15  
Query Match: 69.36% Indels: 1  
DB: 21 Gaps: 0

US-09-786-715-4 (1-118) x AAA31942 (1-419)

OY 4 SerGlucIuGlucIuValIleGlyCysHisSerValAspGluTrpLysGluGlnPheGln 23  
|||||  
DB 44 GCGAGAGGAGGACAGAGTACGCGTGCACACCTGATGCGTGAAGAGCAATTTGAG 103  
OY 24 LysGlyValAspSerLysLysLeuValValIleAspPheThrAlaSerTrpCysGlyPro 43  
|||||  
DB 104 AAGCGGCTCCAAATCCANMAATGCTGTGTGATTTCCACCGCTTCATGTGCGGCCA 163  
OY 44 CysArpPheIleAlaProIleLeuAlaGluMetLysLysThrProHisValIlePhe 63  
|||||  
DB 164 TGCCGCTCTCATTTGCCCAATTTTGTGAAATGGCCAGAAAGATGCCCATGTTTGTTC 223  
OY 64 LeuLysValAspValAspGluLeuLysThrValAlaGluGluPheLysValGluAlaMet 83  
|||||  
DB 224 CTGAGGTGATGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 283  
OY 84 ProThrPheValPheLeuLysGluGlyLysGluValGluArgLeuValGlyAlaArgLys 103  
|||||  
DB 284 CCAACTTTTACTCTTGAAGAGCGCCCAATTT-GTGAGCAANGTGTGGGTGAGACAAA 342  
OY 104 GluGluLeuGlnAlaThrValGluLysHis 113  
|||||  
DB 343 GAACAGTTGCAGGCGGTGATTAACAGACAC 372

RESULT 12  
ABS53097  
ID ABS53097 standard; DNA: 3888 BP.

XX  
XX ABS53097:  
XX  
XX 29-NOV-2002 (first entry)  
XX  
DE DNA encoding Thioredoxin-oleosin fusion protein.  
XX  
XX  
KW Thioredoxin; thioredoxin reductase; gene expression; oleosin;  
KW oil body; oleosin-thioredoxin fusion protein; gene; ds.  
XX  
XX Arabidopsis thaliana.  
XX  
OS Brassica napus.  
XX  
XX Synthetic.  
XX  
XX  
FH Key Location/Qualifiers  
FH CDS 1555..2658  
FT /\*tag= a  
FT /\*product= "Thioredoxin-oleosin fusion protein"  
FT 1555..2250  
FT /\*tag= b  
FT /\*number= 1  
FT intron 2251..2489  
FT /\*tag= c  
FT /\*number= 1  
FT exon 2490..2658  
FT /\*tag= d

```

FT      /number= 2
XX      US2002088025-A1.
XX      04-JUL-2002.
XX      03-JUL-2001; 2001US-0897425.
XX      22-FEB-1991; 91US-0659835.
XX      16-NOV-1993; 93US-0142418.
XX      30-DEC-1994; 94US-0366783.
XX      25-APR-1997; 97US-0846021.
XX      18-DEC-1998; 98US-0210843.
XX      (MOLO/) MOLONEY M. M.
XX      (DALM/) DALMIA B. K.
XX      PI      Moloney MM, Dalmia BK;
XX      WPI: 2002-635723/68.
XX      PT      Expressing protein, by introducing chimeric nucleotide regulatory
XX      PT      sequence, encoding fusion protein, having sequence encoding
XX      PT      protein, oleosin gene and sequence encoding termination region and
XX      PT      producing protein.
XX      Example 21; Fig 14; 69pp; English.
XX      CC      The invention describes a method of expressing thioedoxin or thioedoxin
XX      CC      reductase (I) in the oil body of a host cell using an oil body protein
XX      CC      gene. The method involves introducing a chimeric nucleic acid comprising
XX      CC      a first sequence to regulate transcription, a second DNA sequence
XX      CC      encoding a fusion polypeptide, comprising a sequence encoding an oleosin
XX      CC      gene and sequence encoding (I) and a third sequence encoding a
XX      CC      termination region functional in the host cell and growing the host cell
XX      CC      to produce a fusion polypeptide. The method or (I) is useful for
XX      CC      expression of a thioedoxin or thioedoxin reductase by a host cell. This
XX      CC      sequence encodes a oleosin oil body protein fused to Arabidopsis thaliana
XX      CC      thioedoxin gene controlled by a phaseolin promoter and phaseolin
XX      CC      terminator sequence.
XX      SO      Sequence 3888 BP; 1254 A; 722 C; 613 G; 1299 T; 0 other;

Alignment Scores:
Pred. No.:      5 23e-45      Length:      3888
Score:          420.00      Matches:      74
Percent Similarity: 85.96%      Conservative: 24
Best Local Similarity: 64.91%      Mismatches: 16
Query Match:    69.19%      Indels:      0
DB:             24      Gaps:        0

US-09-786-715-4 (1-118) x ABS53097 (1-3888)
OY      3 SerSerGluGluGlnValIleGlyCysHisSerValAspGluTrpLysGluGlnPhe 22
DB      GCTTCGGAAGAAGAGCAAGTGTGCGCCACACCGTTGACACATGACAGACAGACGCT 1617
OY      23 GlnIysGlyValAspSerLysLysLeuValValIleAspPheThrAlaSerTrpCysGly 42
DB      CAGAAAGGCTAATGATCCAAACACTGTGTGGGTGATTCAACGCGCTTCTGTGGTGA 1677
OY      43 ProCysArgPheIleAlaProIleLeuAlaGluMetAlaLysLysThrProHisValIle 62
DB      CCAATCGTTTCATGCTCCATCTTTCGTGATTTGGCTAAGAAACTTCCTACCTGCTT 1737
OY      63 PheLeuIysValAspValAspGluLeuLysThrValAlaGluGluPheLysValAlaGlu 82
DB      TTCCTCAAGGTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1797
OY      83 MetProThrPheValPheLeuLysGluGluValGluArgLeuValGlyAlaArg 102
DB      ATGCCAACCTTCATGCTTTTGAAGAAGAGGAAATTTGACAAAGTTGTTGAGCAAG 1857

```

```

OY      103 LysGluGluLeuGlnAlaThrValGluLysHisGlyAlaIle 116
DB      AAAAGTACGCTTCACGCTACCATTCGCCAAACACTTGGCTATG 1899
OY      RESULT 13
XX      ABN89581
XX      ID      ABN89581 standard; DNA; 3888 BP.
XX      AC      ABN89581;
XX      DT      06-SEP-2002 (first entry)
XX      DE      Phaseolin promoter-Trxh oleosin-phaseolin terminator DNA SEQ.19.
XX      KW      Multimeric protein; redox protein; thioedoxin; thioedoxin reductase;
XX      KW      oil body; ophthalmological; antidiabetic; cytostatic; antiproliferic;
XX      KW      vasotropic; vulnerrary; antibacterial; immunosuppressive; antitumor;
XX      KW      food product; milk; wheat; oxidative stress; cataract; diabetes;
XX      KW      chronic obstructive pulmonary disease; emphysema; psoriasis; sepsis;
XX      KW      bronchiopulmonary disease; malignancy; reperfusion injury; wound healing;
XX      KW      gastro intestinal bleeding; intestinal bowel disease; ulcer;
XX      KW      gastro oesophageal reflux disease; gene; ds.
XX      OS      Arabidopsis sp.
XX      OS      Phaseolus vulgaris.
XX      PN      WO200250289-A1.
XX      PD      27-JUN-2002.
XX      PF      19-DEC-2001; 2001WO-US50240.
XX      PR      19-DEC-2000; 2000US-0742900.
XX      PR      05-JUL-2001; 2001US-302885P.
XX      PR      04-DEC-2001; 2001US-0006038.
XX      PA      (SEMB-) SEMBIOSYS GENETICS INC.
XX      PA      (SYGN) SYNGENTA PARTICIPATIONS AG.
XX      PI      Van Rooijen G, Deckers H, Helfetz PB, Briggs SP, Dalmia BK;
XX      PI      Del Val G, Zaplachinski S, Moloney M;
XX      DR      WPI: 2002-508806/54.
XX      DR      P-PSDB; ABP60683.
XX      PT      Producing oil body associated with recombinant multimeric protein
XX      PT      complex e.g. redox proteins and immunoglobulins comprises producing
XX      PT      recombinant polypeptides capable of forming the complex in cells
XX      PT      comprising oil bodies.
XX      PS      Example 2; Page 169-171; 362pp; English.
XX      CC      The present invention describes a method (M1) for producing an oil body
XX      CC      associated with a recombinant multimeric protein complex (MPC). M1
XX      CC      comprises producing in a cell comprising oil bodies a first and second
XX      CC      recombinant polypeptide (P1, P2), where P1 is capable of associating
XX      CC      with P2 to form the MPC and associating the complex with an occlusion
XX      CC      body (OB) through an OB-targeting-protein capable of associating with OB
XX      CC      and P1. M1 is useful for producing an oil body associated with a
XX      CC      recombinant MPC. The oil bodies are further formulated for use in the
XX      CC      preparation of a food product such as milk or wheat based food product,
XX      CC      personal care product which reduces the oxidative stress on the surface
XX      CC      area of the human body or used to lighten the skin, or a pharmaceutical
XX      CC      composition used to treat chronic obstructive pulmonary disease (COPD),
XX      CC      catarracts, diabetes, emphysema, bronchiopulmonary disease, psoriasis,
XX      CC      malignancies, reperfusion injury, wound healing, sepsis, gastro
XX      CC      intestinal (GI) bleeding, intestinal bowel disease (IBD), ulcers, GERD
XX      CC      (gastro oesophageal reflux disease). ABN89569 to ABN89593 and ABP60677
XX      CC      to ABP60664 represent sequence given in the exemplification of the
XX      CC      present invention.
XX      SO      Sequence 3888 BP; 1254 A; 722 C; 613 G; 1299 T; 0 other;

```

Alignment Scores: 5.23e-45 Length: 3888  
Pred. No.: 420.00 Matches: 74  
Score: 85.96% Conservative: 24  
Best Local Similarity: 64.91% Mismatches: 16  
Query Match: 69.19% Indels: 0  
DB: 24 Gaps: 0  
US-09-786-715-4 (1-118) x ABN89581 (1-3888)  
QY 3 SerSerGIuGIuGIuValIIleGIyCysHisSerValAspGIuTrpLySGluInPhe 22  
Db 1558 GCTTCGAGAAAGACAGATGATCCCTGCCACACCTTGAGACATGGAACGACACTT 1617  
QY 23 GlnLysGIuValAspSerLysLysLeuValValIIleAspPheThrAlaSerTrpCysGly 42  
Db 1618 CAGAGAGCTAATGAATCAAAACCTTGCTGGTGAATTCACAGGCTTCTGGTGGA 1677  
QY 43 ProCysArgPheIIleAlaProIIleLeuAlaGluMetAlaLysLysThrProHisValIIle 62  
Db 1678 CCATGTCGTTTCATCGCTCCATCTTCTGATTTGGCTAAGAAACTTCTTAACGCTCT 1737  
QY 63 PheLeuLysValAspValAspGIuLeuLysThrValAlaGluInPheLysValGluAla 82  
Db 1738 TTCCTCAAGCTTGATGATGAATTAATCAAGTCGTCAGATGCTTGGCGATACAGCG 1797  
QY 83 MetProThrPheValPheLeuLysGluGlyLysGluValGluArgLeuValGlyAlaArg 102  
Db 1798 ATGCCAACCTTCAATGTTTGAAGAGAGGAAAGATTGGACAAGTTGTTGGAGCCAG 1857  
QY 103 LysGluGluLeuGlnAlaThrValGluLysHisGlyAlaIIle 116  
Db 1858 AAGATGAGCTTCAGTCTACCATTTGCCAAACACTTGGCTATG 1899  
RESULT 14  
ABZ12358 standard; DNA: 345 BP.  
XX ID ABZ12358 standard; DNA: 345 BP.  
XX AC ABZ12358;  
XX XX  
XX 21-JAN-2003 (first entry)  
DE Arabidopsis thaliana stress regulated gene SEQ ID NO 163.  
XX Arabidopsis thaliana; plant; gene; stress; transgenic; ds.  
XX OS  
XX Arabidopsis thaliana.  
XX PN WO200216655-A2.  
XX XX  
XX 28-FEB-2002.  
XX PD  
XX PF 24-AUG-2001; 2001WO-US26685.  
XX XX  
XX PR 24-AUG-2000; 2000US-227866P.  
XX PR 26-JAN-2001; 2001US-26464P.  
XX PR 22-JUN-2001; 2001US-300111P.  
XX XX  
XX PA (SCRI ) SCRIPPS RES-INST.  
XX PA (SYGN ) SYNGENTA PARTICIPATIONS AG.  
XX PI Harper JF, Kreps J, Wang X, Zhu T;  
XX WPI; 2002-304127/34.  
XX DR WPI; 2002-304127/34.  
XX PT Identifying a stress condition to which a plant cell has been exposed  
XX PT and producing plants with increased tolerance to these abiotic stresses  
XX PT  
XX PS Claim 144; SEQ ID NO 163; 577pp + Sequence Listing; English.  
XX CC The invention relates to identifying a stress condition to which a plant  
XX CC cell has been exposed, comprising:

CC (a) contacting nucleic acid representative of expressed polynucleotides  
CC in the plant cell with an array or probes representative of the plant  
CC cell genome; and  
CC (b) detecting a profile of expressed polynucleotides in the plant cell  
CC characteristic of a stress response. The method is useful in the  
CC production of transgenic plants, cells and seeds and in producing plants  
CC with increased tolerance to abiotic stress. The present sequence is that  
CC of an Arabidopsis thaliana stress regulated gene (ABZ12196-ABZ1574) used  
CC in methods of the invention.  
CC Note: The sequence data for this patent is not represented in the printed  
CC specification but is based on sequence information supplied to Derwent by  
CC the European Patent Office.  
XX XX  
XX SQ Sequence 345 BP; 90 A; 68 C; 90 G; 97 T; 0 other;  
Alignment Scores: 2.63e-46 Length: 345  
Pred. No.: 419.00 Matches: 73  
Score: 86.49% Conservative: 23  
Best Local Similarity: 65.77% Mismatches: 15  
Query Match: 69.03% Indels: 0  
DB: 24 Gaps: 0  
US-09-786-715-4 (1-118) x ABZ12358 (1-345)  
QY 3 SerSerGIuGIuGIuValIIleGIyCysHisSerValAspGIuTrpLySGluInPhe 22  
Db 4 GCTTCGAGAAAGACAGATGATCCCTGCCACACCTTGAGACATGGAACGACACTT 63  
QY 23 GlnLysGIuValAspSerLysLysLeuValValIIleAspPheThrAlaSerTrpCysGly 42  
Db 64 CAGAGAGCTAATGAATCAAAACCTTGCTGGTGAATTCACAGGCTTCTGGTGGA 123  
QY 43 ProCysArgPheIIleAlaProIIleLeuAlaGluMetAlaLysLysThrProHisValIIle 62  
Db 124 CCATGTCGTTTCATCGCTCCATCTTCTGATTTGGCTAAGAAACTTCTTAACGCTCT 183  
QY 63 PheLeuLysValAspValAspGIuLeuLysThrValAlaGluInPheLysValGluAla 82  
Db 184 TTCCTCAAGCTTGATGATGAATTAATCAAGTCGTCAGATGCTTGGCGATACAGCG 243  
QY 83 MetProThrPheValPheLeuLysGluGlyLysGluValGluArgLeuValGlyAlaArg 102  
Db 244 ATGCCAACCTTCAATGTTTGAAGAGAGGAAAGATTGGACAAGTTGTTGGAGCCAG 303  
QY 103 LysGluGluLeuGlnAlaThrValGluLysHis 113  
Db 304 AAGATGAGCTTCAGTCTACCATTTGCCAAACAC 336  
RESULT 15  
ABN89587 standard; DNA: 345 BP.  
XX ID ABN89587 standard; DNA: 345 BP.  
XX AC ABN89587;  
XX XX  
XX 06-SEP-2002 (first entry)  
DE Arabidopsis thaliana thioredoxin h (Trx h 1) DNA SEQ ID NO:38.  
XX Arabidopsis thaliana thioredoxin h (Trx h 1) DNA SEQ ID NO:38.  
XX OS  
XX Multimeric protein; redox protein; thioredoxin; thioredoxin reductase;  
XX oil body; ophthalmological; antidiabetic; cytosolic; antipsoriatic;  
XX vasotropic; vulnery; antibacterial; immunosuppressive; antilucer;  
XX food product; milk; wheat; oxidative stress; cataract; diabetes;  
XX chronic obstructive pulmonary disease; emphysema; psoriasis; sepsis;  
XX bronchiolopulmonary disease; malignancy; reperfusion injury; wound healing;  
XX gastro intestinal bleeding; intestinal bowel disease; ulcer;  
XX gastro oesophageal reflux disease; gene; ds.  
XX XX  
XX Arabidopsis thaliana.  
XX OS  
XX PN WO200205289-A1.  
XX XX  
XX 27-JUN-2002.



XX 19-DEC-2001; 2001MO-US50240.  
PF 19-DEC-2000; 2000US-0742900.  
XX 05-JUL-2001; 2001US-302885P.  
PR 04-DEC-2001; 2001US-0006038.  
XX  
PA (SEMB-) SEMBIOSYS GENETICS INC.  
XX (SYGN) SYNGENTA PARTICIPATIONS AG.  
XX Van Rooijen G, Deckers H, Helfetz PB, Briggs SP, Dalmia BK.  
PI Del Val G, Zaplachinski S, Moloney M.  
XX  
DR WPI: 2002-508806/54.  
P-PSDB; ABP60696.  
XX  
PT Producing oil body associated with recombinant multimeric protein  
PT complex e.g. redox proteins and immunoglobulins comprises producing  
PT recombinant polypeptides capable of forming the complex in cells  
PT comprising oil bodies -  
XX  
XX  
PS Claim 68; Page 197; 362pp; English.  
XX  
CC The present invention describes a method (M1) for producing an oil body  
CC associated with a recombinant multimeric protein complex (MPC). M1  
CC comprises producing in a cell comprising oil bodies a first and second  
CC recombinant polypeptide (P1, P2), where P1 is capable of associating  
CC with P2 to form the MPC and associating the complex with an occlusion  
CC body (OB) through an OB-targeting-protein capable of associating with OB  
CC and P1. M1 is useful for producing an oil body associated with a  
CC recombinant MPC. The oil bodies are further formulated for use in the  
CC preparation of a food product such as milk or wheat based food product,  
CC personal care product which reduces the oxidative stress on the surface  
CC area of the human body or used to lighten the skin, or a pharmaceutical  
CC composition used to treat chronic obstructive pulmonary disease (COPD),  
CC cataracts, diabetes, envenomation, bronchiopulmonary disease, psoriasis,  
CC malignancies, reperfusion injury, wound healing, sepsis, gastro  
CC intestinal (GI) bleeding, intestinal bowel disease (IBD), ulcers, GERD  
CC (gastro oesophageal reflux disease). ABN89569 and ABN89593 and ABP60677  
CC to ABP60964 represent sequence given in the exemplification of the  
CC present invention.  
XX  
SQ Sequence 345 BP; 90 A; 68 C; 90 G; 97 T; 0 other;

Alignment Scores:  
Pred. No.: 2.63e-46 Length: 345  
Score: 419.00 Matches: 73  
Percent Similarity: 86.49% Conservative: 23  
Best Local Similarity: 65.77% Mismatches: 15  
Query Match: 69.03% Indels: 0  
DB: 24 Gaps: 0

US-09-786-715-4 (1-118) x ABN89587 (1-345)

OY 3 SerSerGluGluGlnValIleGlyCysHisSerValAspGluThrPylGluGlnPhe 22  
Db 4 GCTTCGAGAGAGAGAAAGTGCCTGCCACACCGCTTGACACATGAGACGACGCTT 63  
OY 23 GlnLysGlyValAspSerLysLysLeuValIleAspPheThrAlaSerTrpCysGly 42  
Db 64 CAGAAAGCCTAATCAACCAAAACTCTTGCTGATTTGGCTAAGAAACCTCTAAGCTGCTT 123  
OY 43 ProGysArgPheIleAlaProIleLeuAlaGluMetAlaLysLysThrProHisValIle 62  
Db 124 CCAATGCTGTTATCGCTCCATCTTGTGCTGATTTGGCTAAGAAACCTCTAAGCTGCTT 183  
OY 63 PheLeuLysValAspValAspGluLeuLysThrValAlaGluLupheLysValGluAla 82  
Db 184 TTCCTCAAGCTGATGATCTGATGATGATGATGATGATGATGATGATGATGATGATG 243  
OY 83 MetProThrPheValPheLeuLysGluGlyValGluValGluArgLeuValGlyAlaArg 102  
Db 244 ATGCCAACCTTATGTTTGAAGAGAGAGAGATTTGGACAAAGTTGTGGAGCCCAAG 303

OY 103 LysGluGluLeuGlnAlaThrValGluLysHis 113  
Db 304 AAAGATGAGCTTCAGTCTACCATTTGCCAAACAC 336

Search completed: August 17, 2003, 19:37:09  
Job time: 167.884 secs



GenCore version 5.1.6  
Copyright (c) 1993 - 2003 CompuGen Ltd.

## OM protein - nucleic search, using frame\_plus\_p2n model

Run on: August 17, 2003, 19:18:57 ; Search time 43.7308 Seconds

(without alignments)  
1190.995 Million cell updates/sec

Title: US-09-786-715-4

Perfect score: 607  
Sequence: 1 MASSEEGVICHSDVEMKE.....VGARKEELQATVEKGAITA 118

## Scoring table:

BLOSUM62	Xgapop 10.0	Xgapext 0.5
Ygapop 10.0	Ygapext 0.5	
Fgapop 6.0	Fgapext 7.0	
Delop 6.0	Delext 7.0	

Searched: 569978 segs, 220691566 residues

Total number of hits satisfying chosen parameters: 1139956

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

## Command line parameters:

-MODEL=frame\_p2n.model -DEV=xlp  
-O=/cgn2\_1/USPTO\_spool/US09786715/runat\_11082003.150515.6083/app\_query.fasta\_1.1052  
-DB=Issued\_Patents.NA -QFMT=fastap -SUFFIX=trn1 -MINMATCH=0.1 -LOOPL=0  
-LOOPEXT=0 -UNITS=DITS -START=1 -END=1 -MATRIX=bl0sum62 -TRANS=human40.cdi  
-LIST=45 -DOCALLIGN=200 -THR.SCORE=pct -THR.MAX=100 -THR.MIN=0 -ALIGN=15  
-MODE=LOCAL -OUTFMT=ptio -NORR=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000  
-USER=US09786715.ecgn1.1.175-@runat\_11082003.150515.6083 -NCPU=6 -ICPU=3  
-NO\_MAP -LARGEQUERY -NEG.SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG  
-DEV.TIMEOUT=120 -WARN.TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6  
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

## Database :

Issued\_Patents.NA:\*  
1: /cgn2\_6/ptodata/1/ina/5A.COMB.seq:\*  
2: /cgn2\_6/ptodata/1/ina/5B.COMB.seq:\*  
3: /cgn2\_6/ptodata/1/ina/6A.COMB.seq:\*  
4: /cgn2\_6/ptodata/1/ina/6B.COMB.seq:\*  
5: /cgn2\_6/ptodata/1/ina/PCUTS.COMB.seq:\*  
6: /cgn2\_6/ptodata/1/ina/backfiles1.seq:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	473.5	78.0	653	1	US-08-181-271A-103
2	473.5	78.0	653	1	US-08-449-315-103
3	473.5	78.0	653	1	US-08-444-803-103
4	473.5	78.0	653	1	US-08-449-043-103
5	473.5	78.0	653	1	US-08-456-265A-103
6	473.5	78.0	653	1	US-08-455-416-103
7	473.5	78.0	653	1	US-08-455-244-103
8	473.5	78.0	653	1	US-08-454-876-103
9	473.5	78.0	653	2	US-08-457-364-103
10	473.5	78.0	653	2	US-08-456-262-103
11	473.5	78.0	653	2	US-08-456-240-103
12	473.5	78.0	653	2	US-08-455-736-103

13	473.5	78.0	653	2	US-08-971-217-103	Sequence 103, App
14	473.5	78.0	653	3	US-09-350-600-103	Sequence 103, App
15	362	59.6	369	4	US-09-540-014-1	Sequence 1, Appl1
16	358	59.0	382	4	US-09-540-014-3	Sequence 3, Appl1
17	358	59.0	393	4	US-09-540-014-5	Sequence 5, Appl1
18	249	41.0	557	4	US-09-404-879A-88	Sequence 88, Appl1
19	249	41.0	557	4	US-09-338-933-88	Sequence 88, Appl1
20	249	41.0	557	4	US-09-215-681-88	Sequence 88, Appl1
21	249	41.0	581	4	US-09-601-144-67	Sequence 67, Appl1
22	249	41.0	594	4	US-09-404-879A-87	Sequence 87, Appl1
23	249	41.0	594	4	US-09-338-933-87	Sequence 87, Appl1
24	249	41.0	601	4	US-09-215-681-87	Sequence 87, Appl1
25	249	41.0	601	4	US-09-404-879A-133	Sequence 133, Appl1
26	249	41.0	601	4	US-09-338-933-133	Sequence 133, Appl1
27	249	41.0	601	4	US-09-215-681-133	Sequence 133, Appl1
28	249	41.0	624	3	US-09-385-982-526	Sequence 526, Appl1
29	249	41.0	630	3	US-08-180-371-5	Sequence 5, Appl1
30	249	41.0	630	3	US-08-180-371-17	Sequence 17, Appl1
31	249	41.0	630	5	PCT-US92-05707-5	Sequence 5, Appl1
32	227.5	37.5	318	4	US-09-313-294A-5868	Sequence 5868, Ap
33	222	36.6	605	3	US-09-385-982-485	Sequence 485, App
34	222	36.6	631	3	US-09-385-982-174	Sequence 174, App
35	211.5	34.8	914	1	US-08-386-729A-9	Sequence 9, Appl1
36	209	34.4	123025	4	US-09-113-294A-5840	Sequence 5840, Ap
37	186	30.6	278	4	US-09-328-352-1234	Sequence 1234, Ap
38	182.5	30.1	438	4	US-08-775-978-2	Sequence 2, Appl1
39	181	29.8	564	2	US-09-103-840A-2	Sequence 1, Appl1
40	179.5	29.6	4403765	3	US-09-103-840A-1	Sequence 1022, Ap
41	179.5	29.6	4411529	3	US-09-134-001C-1022	Sequence 1085, Ap
42	178.5	29.4	336	4	US-09-107-532A-3085	Sequence 373, App
43	172	28.3	339	4	US-09-221-017B-373	Sequence 226, App
44	171	28.2	7096	4	US-08-961-527-226	
45	157.5	25.9	2520	4		

## ALIGNMENTS

RESULT 1  
US-08-181-271A-103  
Sequence 103, Application US/08181271A  
Patent No. 5614395  
GENERAL INFORMATION:  
APPLICANT: Ryals, John A.  
APPLICANT: Alexander, Danny C.  
APPLICANT: Beck, James J.  
APPLICANT: Duesting, John H.  
APPLICANT: Friedlich, Leslie B.  
APPLICANT: Goodman, Robert M.  
APPLICANT: Harms, Christian  
APPLICANT: Meins, Jr., Frederick  
APPLICANT: Montoya, Alice  
APPLICANT: Meyer, Mary B.  
APPLICANT: Neuhaus, Jean-Marc  
APPLICANT: Payne, George B.  
APPLICANT: Sperison, Christoph  
APPLICANT: Stinson, Jeffrey R.  
APPLICANT: Uknes, Scott J.  
APPLICANT: Ward, Eric R.  
TITLE OF INVENTION: WILLIAMS, Sherica C.  
TITLE OF INVENTION: CHEMICALLY REGULATABLE AND ANTI-PATHOGENIC  
TITLE OF INVENTION: DNA SEQUENCES AND USES THEREOF  
NUMBER OF SEQUENCES: 106  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: CIBA-GEIGY Corporation  
STREET: 7 Skyline Drive  
CITY: Hawthorne  
STATE: New York  
COUNTRY: USA  
ZIP: 10532  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/181,271A  
FILING DATE: 13-JAN-94  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/093,301  
FILING DATE: 16-JUL-1993  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/937,197  
FILING DATE: 6-NOV-1992  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/678,378  
FILING DATE: 1-APR-1991  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/305,566  
FILING DATE: 6-FEB-1989  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/165,667  
FILING DATE: 8-MAR-1988  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/042,847  
FILING DATE: 6-APR-1993  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/632,441  
FILING DATE: 21-DEC-1990  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/425,504  
FILING DATE: 20-OCT-1989  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/848,506  
FILING DATE: 6-MAR-1992  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/768,122  
FILING DATE: 27-SEP-1991  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/580,431  
FILING DATE: 7-SEP-1990  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/368,672  
FILING DATE: 20-JUN-1989  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/329,018  
FILING DATE: 24-MAR-1989  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/045,957  
FILING DATE: 12-APR-1993  
ATTORNEY/AGENT INFORMATION:  
NAME: Elmer, James Scott  
REGISTRATION NUMBER: 36,129  
REFERENCE/DOCKET NUMBER: S-19825/PL/COC 1727  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (919)541-8614  
TELEFAX: (919)541-8689  
INFORMATION FOR SEQ ID NO: 103:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 653 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: DNA (genomic)  
US-08-181-271A-103

Alignment Scores:  
Pred. No.: 7.81e-57 Length: 653  
Score: 473.50 Matches: 90  
Percent Similarity: 88.33% Conservative: 16  
Best Local Similarity: 75.00% Mismatches: 10  
Query Match: 78.01% Indels: 4  
DB: 1 Gaps: 1

US-09-786-715-4 (1-118) x US-08-181-271A-103 (1-653)

Qy 3 SerSerGluGluGlnValIleGlyCysHisSerValAspGluTrpLysGluGlnPhe 22  
Db 69 TCATCCGAGGAGGAGCAAGTGTGCGCTGCCACAAAGTTGAGAACTGGAACGACTACTTTC 128  
Qy 23 GlnLysGlyValAspSerLysLysLeuValIleAspPheThrAlaSerTrpCysGly 42  
Db 129 AAGAAAGCGTTGAGACTAAAGAACTGCTGTGCAATTTACTGCTCACTGCTGCGG 188  
Qy 43 ProCysArgPheIleAlaProIleLeuAlaGluMetAlaLysLysThrProHisValIle 62  
Db 189 CCTGCGCGTTTATTTATGCCCATCTCTGCTACACTGTGTAAGAAAGATGCCCATGTTATA 248  
Qy 63 PheLeuLysValAspValAspGluLeuLysThrValAla-GluGluPheLysValGluAl 82  
Db 249 TTCCTCAAGGTGTGATGTGATGAAGTGAAGTCTTTCAGCGGGAATGAGTGTGAGAGC 308  
Qy 82 aMetProThrPheValPheLeuLysGluGlyValGluArgLeuValGlyValAla 102  
Db 309 AATGCCAAGCTTTGCTTCTTCAATTAAGATGAAGAAAGAGTGCACAGAGTGTGCTGCCA 368  
Qy 102 GlySGluGluLeuGlnAlaThrValGluLysHisGlyValA-----IleThrAla 118  
Db 369 GAAGAGAGGTGTGACGACGACCAATAGTGAAGCACTGCTGCTGCTACTGCTACTGCT 426

RESULT 2  
US-08-449-315-103  
Sequence 103, Application US/08449315  
Patent No. 5650505  
GENERAL INFORMATION:  
APPLICANT: Ryals, John A.  
APPLICANT: Alexander, Danny C.  
APPLICANT: Beck, James J.  
APPLICANT: Duesing, John H.  
APPLICANT: Friedrich, Leslie B.  
APPLICANT: Goodman, Robert M.  
APPLICANT: Harms, Christian  
APPLICANT: Meins, Jr., Frederick  
APPLICANT: Montoya, Alice  
APPLICANT: Moyer, Mary B.  
APPLICANT: Neuhaus, Jean-Marc  
APPLICANT: Payne, George B.  
APPLICANT: Sperison, Christoph  
APPLICANT: Stinson, Jeffrey R.  
APPLICANT: Uknes, Scott J.  
APPLICANT: Ward, Eric R.  
APPLICANT: Williams, Shericca C.  
TITLE OF INVENTION: CHEMICALLY REGULATABLE AND ANTI-PATHOGENIC  
NUMBER OF SEQUENCES: 106  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: CIBA-GEIGY Corporation  
STREET: 7 Skyline Drive  
CITY: Hawthorne  
STATE: New York  
COUNTRY: USA  
ZIP: 10532  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/449,315  
FILING DATE: 24-MAY-1995  
CLASSIFICATION: 800  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/181,271  
FILING DATE: 13-JAN-94  
APPLICATION NUMBER: US 08/093,301  
FILING DATE: 16-JUL-1993  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/937,197  
FILING DATE: 6-NOV-1992

```

1  PRIOR APPLICATION DATA:
2  APPLICATION NUMBER: US 07/678,378
3  FILING DATE: 1-APR-1991
4  PRIOR APPLICATION DATA:
5  APPLICATION NUMBER: US 07/305,566
6  FILING DATE: 6-FEB-1989
7  PRIOR APPLICATION DATA:
8  APPLICATION NUMBER: US 07/165,667
9  FILING DATE: 8-MAR-1988
10 PRIOR APPLICATION DATA:
11 APPLICATION NUMBER: US 08/042,847
12 FILING DATE: 6-APR-1993
13 PRIOR APPLICATION DATA:
14 APPLICATION NUMBER: US 07/632,441
15 FILING DATE: 21-DEC-1990
16 PRIOR APPLICATION DATA:
17 APPLICATION NUMBER: US 07/425,504
18 FILING DATE: 20-OCT-1989
19 PRIOR APPLICATION DATA:
20 APPLICATION NUMBER: US 07/848,506
21 FILING DATE: 6-MAR-1992
22 PRIOR APPLICATION DATA:
23 APPLICATION NUMBER: US 07/768,122
24 FILING DATE: 27-SEP-1991
25 PRIOR APPLICATION DATA:
26 APPLICATION NUMBER: US 07/580,431
27 FILING DATE: 7-SEP-1990
28 PRIOR APPLICATION DATA:
29 APPLICATION NUMBER: US 07/368,672
30 FILING DATE: 20-JUN-1989
31 PRIOR APPLICATION DATA:
32 APPLICATION NUMBER: US 07/329,018
33 FILING DATE: 24-MAR-1989
34 PRIOR APPLICATION DATA:
35 APPLICATION NUMBER: US 08/045,957
36 FILING DATE: 12-APR-1993
37 ATTORNEY/AGENT INFORMATION:
38 NAME: Elmer, James Scott
39 REGISTRATION NUMBER: 36,129
40 REFERENCE/DOCKET NUMBER: S-19825/P1/GCG 1727
41 TELECOMMUNICATION INFORMATION:
42 TELEPHONE: (919)541-8614
43 TELEFAX: (919)541-8689
44 INFORMATION FOR SEQ ID NO: 103:
45 SEQUENCE CHARACTERISTICS:
46 LENGTH: 653 base pairs
47 TYPE: nucleic acid
48 STRANDEDNESS: single
49 TOPOLOGY: linear
50 MOLECULE TYPE: DNA (genomic)
51 US-08-449-315-103
52
53 Alignment Scores:
54 Ptd. No.: 7,81e-57 Length: 653
55 Score: 473.50 Matches: 90
56 Percent Similarity: 88.33% Conservative: 16
57 Best Local Similarity: 75.00% Mismatches: 10
58 Query Match: 78.01% Indels: 4
59 DB: 1 Gaps: 1
60
61 US-09-786-715-4 (1-118) x US-08-449-315-103 (1-653)
62
63 3 SerSerGIuGIuGIuValIILegIcySHisSerValaSPGIuTrpLyGIuInPhe 22
64 |||||||  |||||||  |||||||  |||||||  |||||||  |||||||  |||
65 TCTATCGAGAGAGGCAAGTGTTCGGCTGCCCAAGGTTGGAGAGTGAACGAGTACTTC 128
66
67 23 GlnIysGIuValaSPSerLySLysLeuValIILeasPheThrIaSerTrpCysGIy 42
68 ::|||::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
69 AAGAAAGCGCTTGAGACTAAGAACTGGGTGGTGCATTTTACTGCTTATGCTGCGG 188
70
71 43 ProCysArgPheIleAlaTrpIleuAlaGluMetAlaIySLySLTrpProHisValIle 62
72 CATTGGCGGTTTAAATGCCCCCAATTCCTTGTGTGACATTCCTAAGAGATGCCCCCATGTTATA 248
73
74 189 CATTGGCGGTTTAAATGCCCCCAATTCCTTGTGTGACATTCCTAAGAGATGCCCCCATGTTATA 248

```

```

01 66 phenLeuLySValAspValAspGluLeuLysThrValAla-GluGluPheLySValGluAl 82
02 249 TTCTCTAAGAGTGTGATGTTGATGATGAAGACTGTTTCAGCGGGAAGAGGACTGTGGAGGC 308
03 82 aMeCPorThrPheValPheLeuLySGluGlyLySGluValGluArgLeuValGlyAlaAr 102
04 309 AATGCCAACTTTTGTCTTCATCTTAAGAGATGGAAGAAAGATGACAGAGATGTGTGGTCCCA 368
05 Db 369 GAAAGAGAGATTGCAGACAGACCATGATGAAGCATGCTGCTCTACTGTCACACTGCT 426
06
07 RESULT 3
08 Sequence 103, Application US/08444803
09 Patent No. 5654414
10 GENERAL INFORMATION:
11 APPLICANT: Ryals, John A.
12 APPLICANT: Alexander, Danny C.
13 APPLICANT: Beck, James J.
14 APPLICANT: Duesing, John H.
15 APPLICANT: Friedrich, Leslie B.
16 APPLICANT: Goodman, Robert M.
17 APPLICANT: Harms, Christian
18 APPLICANT: Meins, Jr., Frederick
19 APPLICANT: Montoya, Alice
20 APPLICANT: Moyer, Mary B.
21 APPLICANT: Neuhaus, Jean-Marc
22 APPLICANT: Payne, George B.
23 APPLICANT: Sperison, Christoph
24 APPLICANT: Stinson, Jeffrey R.
25 APPLICANT: Uknes, Scott J.
26 APPLICANT: Ward, Eric R.
27 APPLICANT: Williams, Shericca C.
28 TITLE OF INVENTION: CHEMICALLY REGULATABLE AND ANTI-PATHOGENIC
29 TITLE OF INVENTION: DNA SEQUENCES AND USES THEREOF
30 NUMBER OF SEQUENCES: 106
31 CORRESPONDENCE ADDRESS:
32 ADDRESSEE: CIBA-GEIGY Corporation
33 STREET: 7 Skyline Drive
34 CITY: Hawthorne
35 STATE: New York
36 COUNTRY: USA
37 ZIP: 10532
38
39 COMPUTER READABLE FORM:
40 MEDIUM TYPE: Floppy disk
41 COMPUTER: IBM PC compatible
42 OPERATING SYSTEM: PC-DOS/MS-DOS
43 SOFTWARE: Patentin Release #1.0, Version #1.25
44 CURRENT APPLICATION DATA:
45 APPLICATION NUMBER: US/08/444,803
46 FILING DATE: 19-MAY-1995
47 CLASSIFICATION: 536
48 PRIOR APPLICATION DATA:
49 APPLICATION NUMBER: 08/181,271
50 FILING DATE: 13-JAN-94
51 APPLICATION NUMBER: US 08/093,301
52 FILING DATE: 16-JUL-1993
53 PRIOR APPLICATION DATA:
54 APPLICATION NUMBER: US 07/937,197
55 FILING DATE: 6-NOV-1992
56 PRIOR APPLICATION DATA:
57 APPLICATION NUMBER: US 07/678,378
58 FILING DATE: 1-APR-1991
59 PRIOR APPLICATION DATA:
60 APPLICATION NUMBER: US 07/305,566
61 FILING DATE: 6-FEB-1989
62 PRIOR APPLICATION DATA:
63 APPLICATION NUMBER: US 07/165,667
64 FILING DATE: 8-MAR-1988
65 PRIOR APPLICATION DATA:
66 APPLICATION NUMBER: US 08/042,847

```



APPLICATION NUMBER: US 07/768,122  
FILING DATE: 27-SEP-1991  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/580,431  
FILING DATE: 7-SEP-1990  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/368,672  
FILING DATE: 20-JUN-1989  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/329,018  
FILING DATE: 24-MAR-1989  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/045,957  
FILING DATE: 12-APR-1993  
ATTORNEY/AGENT INFORMATION:  
NAME: Elmer, James Scott  
REGISTRATION NUMBER: 36,129  
REFERENCE/DOCKET NUMBER: S-19825/PL/CGC 1727  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (919)541-8614  
TELEFAX: (919)541-8689  
INFORMATION FOR SEQ ID NO: 103:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 653 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: DNA (genomic)  
US-08-449-043-103

Alignment Scores:  
Pred. No.: 7,81e-57 Length: 653  
Score: 473.50 Matches: 90  
Percent Similarity: 88.33% Conservative: 16  
Best Local Similarity: 75.00% Mismatches: 10  
Query Match: 78.01% Indels: 4  
Gaps: 1

US-09-786-715-4 (1-118) x US-08-449-043-103 (1-653)

OY 3 SerSerGluGluGluGluValIleGlyCysHisSerValAspGluTrpLysGluInphe 22  
Db 69 TCATCCGAGGAGGAGGACAGTGTTCGCGCCACAGAGTTGAGAGTGAAGACGATCTTC 128  
OY 23 GlnYssGlyValAspSerLysLysLeuValIleAspPheThrAlaSerTrpYsgly 42  
Db 129 AAGAAAGCGGTGAAGTAAGAACTGGTGGTGGATTTACTGCTTCATGAGTGGCGS 188  
OY 43 ProCysArgPheIleAlaProIleLeuAlaGluMeTAlaLysLysThrProHisValIle 62  
Db 189 CCTTCCGCTTTATTTGCCCCCAATTTCTTGTGCTGATGCTAGAGAAATGCCCATTTATA 248  
OY 63 PheLeuLysValAspValAspGluLeuLysThrValAla-GluGluPheLysValGluAl 82  
Db 249 TTCCCAAGGTGTGATGTGATGAAGACTGTTTCAGCGGGAATGAGTGTGAGAGC 308  
OY 82 aMePProhPheValPheLeuLysGluGlyLysValAlaGluArgLeuValGlyAlaAr 102  
Db 309 AATGCCAATTTGCTTCTCATTAAGATGAAGAAAGTGAAGAGTGGTGGGCCAA 368  
OY 102 gLysGluGluLeuGlnAlaThrValGluLysHisGlyAla-----IleThrAla 118  
Db 369 GAAGAGAGAGTTGCACGACGACACCATATGTGAAGATCTCTCTCTACTGTCTACTGCT 426

RESULT 5

US-08-456-265A-103  
Sequence 103, Application US/08456265A  
Patent No. 5767369  
GENERAL INFORMATION:  
APPLICANT: Alexander, Danny C.  
APPLICANT: Ryals, John A.  
APPLICANT: Goodman, Robert M.  
APPLICANT: Stinson, Jeffrey R.

TITLE OF INVENTION: CHEMICALLY REGULATABLE AND ANTI-PATHOGENIC  
TITLE OF INVENTION: DNA SEQUENCES AND USES THEREOF  
NUMBER OF SEQUENCES: 111  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: CIBA-GEIGY Corporation  
STREET: 520 White Plains Road, P.O. Box 2005  
CITY: Tarrytown  
STATE: New York  
COUNTRY: USA  
ZIP: 10591  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/456,265A  
FILING DATE: 31-MAY-95  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/181,271  
FILING DATE: 13-JAN-1994  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/093,301  
FILING DATE: 16-JUL-1993  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/937,197  
FILING DATE: 6-NOV-1992  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/678,378  
FILING DATE: 1-APR-1991  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/305,566  
FILING DATE: 6-FEB-1989  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/165,667  
FILING DATE: 8-MAR-1988  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/042,847  
FILING DATE: 6-APR-1993  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/632,441  
FILING DATE: 21-DEC-1990  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/425,504  
FILING DATE: 20-OCT-1989  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/848,506  
FILING DATE: 6-MAR-1992  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/768,122  
FILING DATE: 27-SEP-1991  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/580,431  
FILING DATE: 7-SEP-1990  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/368,672  
FILING DATE: 20-JUN-1989  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/329,018  
FILING DATE: 24-MAR-1989  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/045,957  
FILING DATE: 12-APR-1993  
ATTORNEY/AGENT INFORMATION:  
NAME: Meigs, J. Timothy  
REGISTRATION NUMBER: 38,241  
REFERENCE/DOCKET NUMBER: S-19825/PL/CGC 1727/DIV10  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (919)541-8587  
TELEFAX: (919)541-8689  
INFORMATION FOR SEQ ID NO: 103:  
SEQUENCE CHARACTERISTICS:

```

;      LENGTH: 653 base pairs
;      TYPE: nucleic acid
;      STRANDEDNESS: single
;      TOPOLOGY: linear
;      MOLECULE TYPE: DNA (genomic)
US-08-456-265A-103

```

Alignment Scores:	
Pred. No.:	7.81e-57
Score:	473.50
Percent Similarity:	88.33%
Best Local Similarity:	75.00%
Query Match:	78.01%
DB:	1
Length:	655
Matches:	90
Conservative:	16
Mismatches:	10
Indels:	4
Gaps:	1

US-09-786-715-4 (1-118) x US-08-456-265A-103 (1-653)

```

QY      serSeGcIuGlGcIvAlIleGcYshSerValAspGluTrpYsGIuGlnPhe  22
Db      69  TCATCGAGGAGGCAACAGTGTTCGGCTGCCCAAGGTGTGAGAAATGCAACGACTTTC  122
QY      23  GlnYsGcIYValAspSerLysLysLeuValIleAspPheThrAlaSerTrpCysGly  42
Db      129  AAGAAAGGGCGTTGACCTAGAAACCTGGCGGGCGATTTCACGCTTCATCGTGGCGGS  188
QY      43  ProCysArgPheIleAlaProIleuAlaGluMetAlaLysTrpProHisValIle  62
Db      189  CTTTCGCGTTTATTGCCCAATCTTGCTGACATGTCTAAGAGATGCCCAATGTTATA  248
QY      63  PheLeuLysValAspValAspGluLeuLysThrValAlaL-GluGluPheLysValGluAl  82
Db      249  TTCTCTAACCTTGATCTTGATGATCACTGAGGACTGTTTCAGCGGGAATGCACTGGGAGGC  308
QY      82  aMetProThrPheValPheLeuLysGcIuGlyYsGcIuValGluArgLeuValGlyAlaAr  102
Db      309  AATGCCAATCTTTGCTTCATTAAAGATGGAAGAAAGAAAGATGACAGCTTGTTGGTCCAA  368
QY      102  gYsGcIuGluGlnAlaThrValGcIuLysHisGlyAla-----IleThrAla  118
Db      369  GAAAGGAGGATGGCACAGCAATGATGAGATGCTCTCTCGTACTGCACTGGTGT  426

```

RESULT 6  
 US-08-455-416-103  
 : Sequence 103, Application US/08455416  
 Patent No. 5777200  
 GENERAL INFORMATION:  
 APPLICANT: Ryals, John A.  
 APPLICANT: Alexander, Danny C.  
 APPLICANT: Beck, James J.  
 APPLICANT: Duesing, John H.  
 APPLICANT: Friedrich, Leslie B.  
 APPLICANT: Goodman, Robert M.  
 APPLICANT: Harms, Christian  
 APPLICANT: Melns, Jr., Frederick  
 APPLICANT: Montoya, Alice  
 APPLICANT: Moyer, Mary B.  
 APPLICANT: Neubaus, Jean-Marc  
 APPLICANT: Payne, George B.  
 APPLICANT: Sperlison, Christoph  
 APPLICANT: Stinson, Jeffrey R.  
 APPLICANT: Uknes, Scott J.  
 APPLICANT: Ward, Eric R.  
 APPLICANT: Williams, Sherlica C.  
 TITLE OF INVENTION: CHEMICALS THAT REGULATE AND ANTI-PATHOGENIC  
 TITLE OF INVENTION: DNA SEQUENCES AND USES THEREOF  
 NUMBER OF SEQUENCES: 106  
 CORRESPONDENCE ADDRESS:

```

:      COMPUTER READABLE FORM:
:      MEDIUM TYPE: Floppy disk
:      COMPUTER: IBM PC compatible
:      OPERATING SYSTEM: PC-DOS/MS-DOS
:      SOFTWARE: PatentIn Release #1.0, Version #1.25
:      CURRENT APPLICATION DATA:

```

CITY:	HAWTHORNE	Alignment Scores:	
STATE:	New York	Pred. No.:	7.81e-57
COUNTRY:	USA	Score:	473.50
ZIP:	10532	Percent Similarity:	88.33%
		Length:	655
		Matches:	90
		Conservative:	16



```
Best Local Similarity: 75.00% Mismatches: 10
Query Match: 78.01% Indels: 4
DB: 1 Gaps: 1
US-09-786-715-4 (1-118) x US-08-455-416-103 (1-653)
OY 3 SerSerGIUGLUGIYGLInValIleGlyCysHISerValAspGLUtrPrLysGLuInPhe 22
   |||||
DB 69 TCATCCGAGGAGGACAAAGTGTGGCTCCCAAGGTTGAGAGTAAGCAAGTACTTCTTC 128
OY 23 GlnLysGlyValAspSerLysLysLeuValIleAspPheThrAlaSerTrpCysGly 42
   |||||
DB 129 AAGAAAGCGCTGAGACTTAAGAAAGTGTGGTGTGATTTACTGCTTCATGAGTCCGGS 188
OY 43 ProCysArgPheIleAlaProIleLeuAlaGluMetAlaLysLysThrProHISValIle 62
   |||||
DB 189 CTTTCCTCCGTTTATTTGCCCCCAATTCCTGTGACATTCCTAAAGATGCCCCATGTTATA 248
OY 63 PheLeuLysValAspValAspLysLysLeuValIleGluLysPheLysValGluAla 82
   |||||
DB 249 TTCCTCAAGGTGATGATTCATGACAGCAAGACTGTTTCACGGGAATGAGTGTGAGGC 308
OY 82 aMetProThrPheValIlePheLeuLysGluGlyLysGluValIleGluArgLeuValGlyAlaAr 102
   |||||
DB 309 AATGCCAAGTCTTCTTCATTAAGATGAAAGAGTGAAGAGTGTGTTGTTGCCAA 368
OY 102 GlySGluGluLeuGlnAlaThrValGluLysHISGlyAla-----IleThrAla 118
   |||||
DB 369 GAAGAAGAGTTCACACAGACCATATGATGAGATGCTGCTCTACTGACTACTGCT 426

RESULT 7
US-08-455-244-103
: Sequence 103, Application US/08455244
: Patent No 5789214
: GENERAL INFORMATION:
: APPLICANT: Ryals, John A.
: APPLICANT: Alexander, Danny C.
: APPLICANT: Beck, James J.
: APPLICANT: Duesing, John H.
: APPLICANT: Friedrich, Leslie B.
: APPLICANT: Goodman, Robert M.
: APPLICANT: Harms, Christian
: APPLICANT: Meins, Jr., Frederick
: APPLICANT: Montoya, Alice
: APPLICANT: Moyer, Mary B.
: APPLICANT: Neuhaus, Jean-Marc
: APPLICANT: Payne, George B.
: APPLICANT: Sperison, Christoph
: APPLICANT: Stinson, Jeffrey R.
: APPLICANT: Uknes, Scott J.
: APPLICANT: Ward, Eric R.
: APPLICANT: Williams, Shericea C.
: TITLE OF INVENTION: CHEMICALLY REGULATABLE AND ANTI-PATHOGENIC
: NUMBER OF SEQUENCES: 106
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: CIBA-GEIGY Corporation
: STREET: 7 Skyline Drive
: CITY: Hawthorne
: STATE: New York
: COUNTRY: USA
: ZIP: 10532
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: PatentIn Release #1.0, Version #1.25
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/08/455,244
: FILING DATE: 31-MAY-1995
: CLASSIFICATION: 435
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: 08/181,271
```

```

Db      129 AAGAAAGCGCTTGAGACTAGAAAGCTGCTGCTCATTTTACTGCTTTCATGCTGCGGS 188
      43 ProcysarphetheleaproleleuAlaglumetalalyslsthProhisValIle 62
      189 CTTGGCCGTTTATTTGCCCCCAATTTCTTGCTGACATTTCTAGAGAAGATGCCCATGTTATA 248
Qy      63 PheleulysValaspValaspGluLeulysThrValAla-GluGluPheulysValGluAl 82
      249 TTCCTCAAGGTTGATGTGATGAACCTGTTTCACGCGGGAATGGAGTGTGGAGGC 308
Qy      82 aMetProthPheValPheleulysGluGlyLysGluValGluArgLeuValGlyAlaAr 102
      309 AATGCCAAGCTTTGCTCTTCATTTAAAGATGGAAAGAGTGGACAGCTTGTGTCGCAA 368
Qy      102 gLysGluLeuLeuAlaThrValGluLysHisGlyVala-----IleThrAla 118
      369 GAAAGAGAGCTTGACAGCAGACCATAGTGAAGCATGCTGCTCTGCTACTGTCAGTGT 426

RESULT 8
US-08-454-876-103
; Sequence 103, Application US/08454876
; Patent No. 5804693
; GENERAL INFORMATION:
; APPLICANT: Ryals, John A.
; APPLICANT: Alexander, Danny C.
; APPLICANT: Beck, James J.
; APPLICANT: Duesing, John H.
; APPLICANT: Friedrich, Leslie B.
; APPLICANT: Goodman, Robert M.
; APPLICANT: Harms, Christian
; APPLICANT: Helms, Jr., Frederick
; APPLICANT: Montoya, Alice
; APPLICANT: Moyer, Mary B.
; APPLICANT: Neuhaus, Jean-Marc
; APPLICANT: Payne, George B.
; APPLICANT: Sperison, Christoph
; APPLICANT: Stinson, Jeffrey R.
; APPLICANT: Uknes, Scott J.
; APPLICANT: Ward, Eric R.
; APPLICANT: Williams, Sherica C.
; TITLE OF INVENTION: CHEMICALLY REGULATABLE AND ANTI-PATHOGENIC
; NUMBER OF SEQUENCES: 106
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: CIBA-GEIGY Corporation
; STREET: 7 Skyline Drive
; CITY: Hawthorne
; STATE: New York
; COUNTRY: USA
; ZIP: 10532
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/454,876
; FILING DATE: 31-MAY-1995
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/181,271
; FILING DATE: 13-JAN-94
; APPLICATION NUMBER: US 08/093,301
; FILING DATE: 16-JUL-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/937,197
; FILING DATE: 6-NOV-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/678,378
; FILING DATE: 1-APR-1991
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/305,566

```

```

; FILING DATE: 6-FEB-1989
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/165,667
; FILING DATE: 8-MAR-1988
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/042,847
; FILING DATE: 6-APR-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/632,441
; FILING DATE: 21-DEC-1990
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/425,504
; FILING DATE: 20-OCT-1989
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/848,506
; FILING DATE: 6-MAR-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/768,122
; FILING DATE: 27-SEP-1991
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/580,431
; FILING DATE: 7-SEP-1990
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/368,672
; FILING DATE: 20-JUN-1989
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/329,018
; FILING DATE: 24-MAR-1989
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/045,957
; FILING DATE: 12-APR-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: Elmer, James Scott
; REGISTRATION NUMBER: 36,129
; REFERENCE/DOCKET NUMBER: S-19825/PL/CGC 1727
; TELEPHONE: (919)541-8614
; TELEFAX: (919)541-8689
; INFORMATION FOR SEQ ID NO: 103:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 653 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; US-08-454-876-103

Alignment Scores:
Pred. No.: 7,81e-57 Length: 653
Score: 473.50 Matches: 90
Percent Similarity: 88.33% Conservative: 16
Best Local Similarity: 75.00% Mismatches: 10
Query Match: 78.01% Indels: 4
DB: 1 Gaps: 1

US-09-786-715-4 (1-118) x US-08-454-876-103 (1-653)
Qy      3 SerSerGluGluGlyGlnValIleGlyCysHisSerValaspGluThrPlyGluGlnPhe 22
      69 TCATCCGAGAGGAGACAGAGTTCGCGTCCGACAGAGTTGAGGAATGGAAGAGTACTTC 128
Db      129 AAGAAAGCGCTTGAGACTAGAAAGCTGCTGCTCATTTTACTGCTTTCATGCTGCGGS 188
Qy      23 GlnLysGlyValaspSerLysLysLeuValValIleAspPheThrAlaSerTPCyGly 42
      43 ProcysarphetheleaproleleuAlaglumetalalyslsthProhisValIle 62
Db      189 CTTGGCCGTTTATTTGCCCCCAATTTCTTGCTGACATTTCTAGAGAAGATGCCCATGTTATA 248
Qy      63 PheleulysValaspValaspGluLeulysThrValAla-GluGluPheulysValGluAl 82
      249 TTCCTCAAGGTTGATGTGATGAACCTGTTTCACGCGGGAATGGAGTGTGGAGGC 308

```

Qy 82 ametProthrpheValPheLeuLysGluGlyValGluArgLeuValGlyAlaAr 102  
Db 309 AATGCCAACTTTTCTTCATTAAGATGAAAGAACTGACAGATGTTGGTCCAA 368  
Qy 102 glysGluGluLeuGlnAlaThrValGluLysHisGlyAla-----IleThrAla 118  
Db 369 GAAGAGGAGTTGCAGCAGACCAATAGTGAGCAATGCTCTGCTACTGTACACTGCT 426

RESULT 9  
US-08-457-364-103  
Sequence 103, Application US/08457364  
Patent No. 5847258  
GENERAL INFORMATION:  
APPLICANT: Ryals, John A.  
APPLICANT: Alexander, Danny C.  
APPLICANT: Beck, James J.  
APPLICANT: Duesting, John H.  
APPLICANT: Friedlich, Leslie B.  
APPLICANT: Goodman, Robert M.  
APPLICANT: Harms, Christian  
APPLICANT: Meins, Jr., Frederick  
APPLICANT: Montoya, Alice  
APPLICANT: Moyer, Mary B.  
APPLICANT: Neuhaus, Jean-Marc  
APPLICANT: Payne, George B.  
APPLICANT: Sperison, Christoph  
APPLICANT: Stinson, Jeffrey R.  
APPLICANT: Uknes, Scott J.  
APPLICANT: Ward, Eric R.  
APPLICANT: Williams, Shericea C.  
TITLE OF INVENTION: CHEMICALLY REGULATABLE AND ANTI-PATHOGENIC  
TITLE OF INVENTION: DNA SEQUENCES AND USES THEREOF  
NUMBER OF SEQUENCES: 106  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: CIBA-GEIGY Corporation  
STREET: 7 Skyline Drive  
CITY: Hawthorne  
STATE: New York  
COUNTRY: USA  
ZIP: 10532  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/457,364  
FILING DATE: 31-MAY-1995  
CLASSIFICATION: 800  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/181,271  
FILING DATE: 13-JAN-94  
APPLICATION NUMBER: US 08/093,301  
FILING DATE: 16-JUL-1993  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/937,197  
FILING DATE: 6-NOV-1992  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/678,378  
FILING DATE: 1-APR-1991  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/305,566  
FILING DATE: 6-FEB-1989  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/165,667  
FILING DATE: 8-MAR-1988  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/042,847  
FILING DATE: 6-APR-1993  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/632,441  
FILING DATE: 21-DEC-1990  
PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 07/425,504  
FILING DATE: 20-OCT 1989  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/848,506  
FILING DATE: 6-MAR-1992  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/768,122  
FILING DATE: 27-SEP-1991  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/580,431  
FILING DATE: 7-SEP-1990  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/368,672  
FILING DATE: 20-JUN-1989  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/329,018  
FILING DATE: 24-MAR-1989  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/045,957  
FILING DATE: 12-APR-1993  
ATTORNEY/AGENT INFORMATION:  
NAME: Elmer, James Scott  
REGISTRATION NUMBER: 36,129  
REFERENCE/DOCKET NUMBER: S-19825/P1/CGC 1727  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (919)541-8614  
TELEFAX: (919)541-8689  
INFORMATION FOR SEQ ID NO: 103:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 653 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: DNA (genomic)  
US-08-457-364-103

Alignment Scores:  
Pred. No.: 7,81e-57 Length: 653  
Score: 473.50 Matches: 90  
Percent Similarity: 88.33% Conservative: 16  
Best Local Similarity: 75.00% Mismatches: 10  
Query Match: 78.01% Indels: 4  
DB: 2 Gaps: 1

US-09-786-715-4 (1-118) x US-08-457-364-103 (1-653)

Qy 3 SerSergLugluGluValIleGlyCysHisSerValAspGluTrpLysGluGlnPhe 22  
Db 69 TCATCCGAGGAGGAGCAGATGTTGGCTGCCACAGAGTTGAGATGAGACGATGACTTC 128  
Qy 23 GlnLysGlyValAspSerLysLysLeuValIleAspPheThrAlaSerTrpCysGly 42  
Db 129 AAGAAAGCGTTGAGACATGAGAACTGCTGGTGGTCACTTACTGCTTCAATGGCGGS 188  
Qy 43 ProCysArgPheIleAlaProIleLeuAlaGluMetAlaLysThrProHisValIle 62  
Db 189 CCTTGCCTTTTATTTGCCCAATTTCTTCTGACATGCTTAAGACATGCCCATGTTTATA 248  
Qy 63 PheLeuLysValAspValAspGluLeuLysThrValAla-GluGluPheLysValGluAl 82  
Db 249 TTCTCCAGAGTTGATGATGATGAGACTGAAGACTGTTTCAGCGGGAATGAGAGTGGAGGC 308  
Qy 82 ametProthrpheValPheLeuLysGluGlyValGluArgLeuValGlyAlaAr 102  
Db 309 AATGCCAACTTTTCTTCATTAAGATGAAAGAACTGACAGATGTTGGTCCAA 368  
Qy 102 glysGluGluLeuGlnAlaThrValGluLysHisGlyAla-----IleThrAla 118  
Db 369 GAAGAGGAGTTGCAGCAGACCAATAGTGAGCAATGCTCTGCTACTGTACACTGCT 426

RESULT 10  
US-08-456-262-103  
Sequence 103, Application US/08456262



APPLICANT: Moyer, Mary B.  
APPLICANT: Neuhaus, Jean-Marc  
APPLICANT: Payne, George B.  
APPLICANT: Sperison, Christoph  
APPLICANT: Stinson, Jeffrey R.  
APPLICANT: Utnes, Scott J.  
APPLICANT: Ward, Eric R.  
APPLICANT: Williams, Shericca C.  
TITLE OF INVENTION: CHEMICALLY REGULATABLE AND ANTI-PATHOGENIC  
TITLE OF INVENTION: DNA SEQUENCES AND USES THEREOF  
NUMBER OF SEQUENCES: 106  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: CIBA-GEIGY Corporation  
STREET: 7 Skyline Drive  
CITY: Hawthorne  
STATE: New York  
COUNTRY: USA  
ZIP: 10532  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/456,240  
FILING DATE: 31-MAY-1995  
CLASSIFICATION: 800  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/181,271  
FILING DATE: 13-JAN-94  
APPLICATION NUMBER: US 08/093,301  
FILING DATE: 16-JUL-1993  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/937,197  
FILING DATE: 6-NOV-1992  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/678,378  
FILING DATE: 1-APR-1991  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/305,566  
FILING DATE: 6-FEB-1989  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/165,667  
FILING DATE: 8-MAR-1988  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/042,847  
FILING DATE: 6-APR-1993  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/632,441  
FILING DATE: 21-DEC-1990  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/425,504  
FILING DATE: 20-OCT-1989  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/848,506  
FILING DATE: 6-MAR-1992  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/768,122  
FILING DATE: 27-SEP-1991  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/580,431  
FILING DATE: 7-SEP-1990  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/368,672  
FILING DATE: 20-JUN-1989  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/329,018  
FILING DATE: 24-MAR-1989  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/045,957  
FILING DATE: 12-APR-1993  
ATTORNEY/AGENT INFORMATION:  
NAME: Elmer, James Scott

REGISTRATION NUMBER: 36,129  
REFERENCE/DOCKET NUMBER: S-19825/P1/CSC 1727  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (919)541-8614  
TELEFAX: (919)541-8689  
INFORMATION FOR SEQ ID NO: 103:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 653 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: DNA (genomic)  
US-08-456-240-103  
Alignment Scores:  
Pred. No.: 7,816-57  
Score: 473.50  
Percent Similarity: 88.33%  
Best Local Similarity: 75.00%  
Query Match: 78.01%  
DB: 2  
Gaps: 1  
US-09-786-715-4 (1-118) x US-08-456-240-103 (1-653)  
QY 3 SerSerGluGluGluValIleGlycSHisSerValAspGluTrpLysGluInpHe 22  
DB 69 TCATCCGAGGAGGACAAAGTGTGCGCCCAAGGTTGAGAGATGAGACGATCTTC 128  
QY 23 GlnLysGlyValAspSerLysLysLeuValIleAspPheThAlaSerTrpCysGly 42  
DB 129 AAGAAAGCGCTTGACACTAAGAAACTGGTGGTGCATTTTACGCTTCAAGCGCGS 188  
QY 43 ProCysArgPheIleAlaProIleLeuAlaGluMetAlaLysLysTrpHisValIle 62  
DB 189 CCTTCGCTTTATTTGCCCCCAATCTCTGACATTGCTAAGAAAGATGCGCATGTATA 248  
QY 63 PheLeuLysValAspValAspGluLeuLysThrValAla-GluGluPheLysValGluAl 82  
DB 249 TTCCTCAAGGTGATGATGATGACTGAAGACTGTTTCAGCGGATGAGTGTGAGGC 308  
QY 82 aMePProThrPheValPheLeuLysGluGlyLysGluValGluArgLeuValGlyAla 102  
DB 309 AATGCCACTTTTGTCTTCATTAAAGATGAAAGAAAGTGAACAGATGTTGTCCTCA 368  
QY 102 GlySGluGluLeuGlnAlaThrValGluLysHisGlyAla-----IleThrAla 118  
DB 369 GAAAGAGGAGTTGCAGACACCATAGTGAAGCATGCTCCTCTACTGTCATGCT 426  
RESULT 12  
US-08-456-736-103  
Sequence 103, Application US/08455736  
Patent No. 5880328  
GENERAL INFORMATION:  
APPLICANT: Ryals, John A.  
APPLICANT: Alexander, Danny C.  
APPLICANT: Beck, James J.  
APPLICANT: Duesing, John H.  
APPLICANT: Friedrich, Leslie B.  
APPLICANT: Goodman, Robert M.  
APPLICANT: Harms, Christian  
APPLICANT: Meins, Jr., Frederick  
APPLICANT: Montoya, Alice  
APPLICANT: Moyer, Mary B.  
APPLICANT: Neuhaus, Jean-Marc  
APPLICANT: Payne, George B.  
APPLICANT: Sperison, Christoph  
APPLICANT: Stinson, Jeffrey R.  
APPLICANT: Utnes, Scott J.  
APPLICANT: Ward, Eric R.  
APPLICANT: Williams, Shericca C.  
TITLE OF INVENTION: CHEMICALLY REGULATABLE AND ANTI-PATHOGENIC  
TITLE OF INVENTION: DNA SEQUENCES AND USES THEREOF  
NUMBER OF SEQUENCES: 106

;; CORRESPONDENCE ADDRESS:  
;; ADDRESSEE: CIBA-GEIGY Corporation  
;; STREET: 7 Skyline Drive  
;; CITY: Hawthorne  
;; STATE: New York  
;; COUNTRY: USA  
;; ZIP: 10532  
;;  
;; COMPUTER READABLE FORM:  
;; MEDIUM TYPE: Floppy disk  
;; COMPUTER: IBM PC compatible  
;; OPERATING SYSTEM: PC-DOS/MS-DOS  
;; SOFTWARE: Patentin Release #1.0, Version #1.25  
;;  
;; CURRENT APPLICATION DATA:  
;; APPLICATION NUMBER: US/08/455,736  
;; FILING DATE: 31-MAY-1995  
;; CLASSIFICATION: 435  
;;  
;; PRIOR APPLICATION DATA:  
;; APPLICATION NUMBER: 08/181,271  
;; FILING DATE: 13-JAN-1994  
;; APPLICATION NUMBER: US 08/093,301  
;; FILING DATE: 16-JUL-1993  
;;  
;; PRIOR APPLICATION DATA:  
;; APPLICATION NUMBER: US 07/937,197  
;; FILING DATE: 6-NOV-1992  
;;  
;; PRIOR APPLICATION DATA:  
;; APPLICATION NUMBER: US 07/678,378  
;; FILING DATE: 1-APR-1991  
;;  
;; PRIOR APPLICATION DATA:  
;; APPLICATION NUMBER: US 07/305,566  
;; FILING DATE: 6-FEB-1989  
;;  
;; PRIOR APPLICATION DATA:  
;; APPLICATION NUMBER: US 07/165,667  
;; FILING DATE: 8-MAR-1988  
;;  
;; PRIOR APPLICATION DATA:  
;; APPLICATION NUMBER: US 08/042,847  
;; FILING DATE: 6-APR-1993  
;;  
;; PRIOR APPLICATION DATA:  
;; APPLICATION NUMBER: US 07/632,441  
;; FILING DATE: 21-DEC-1990  
;;  
;; PRIOR APPLICATION DATA:  
;; APPLICATION NUMBER: US 07/425,504  
;; FILING DATE: 20-OCT-1989  
;;  
;; PRIOR APPLICATION DATA:  
;; APPLICATION NUMBER: US 07/848,506  
;; FILING DATE: 6-MAR-1992  
;;  
;; PRIOR APPLICATION DATA:  
;; APPLICATION NUMBER: US 07/768,122  
;; FILING DATE: 27-SEP-1991  
;;  
;; PRIOR APPLICATION DATA:  
;; APPLICATION NUMBER: US 07/580,431  
;; FILING DATE: 7-SEP-1990  
;;  
;; PRIOR APPLICATION DATA:  
;; APPLICATION NUMBER: US 07/368,672  
;; FILING DATE: 20-JUN-1989  
;;  
;; PRIOR APPLICATION DATA:  
;; APPLICATION NUMBER: US 07/329,018  
;; FILING DATE: 24-MAR-1989  
;;  
;; PRIOR APPLICATION DATA:  
;; APPLICATION NUMBER: US 08/045,957  
;; FILING DATE: 12-APR-1993  
;;  
;; ATTORNEY/AGENT INFORMATION:  
;; NAME: Elmer, James Scott  
;; REGISTRATION NUMBER: 36,129  
;; REFERENCE/DOCKET NUMBER: S-19825/PL/CGC 1727  
;; TELECOMMUNICATION INFORMATION:  
;; TELEPHONE: (919)541-8614  
;; TELEFAX: (919)541-8689  
;;  
;; INFORMATION FOR SEQ ID NO: 103:  
;; SEQUENCE CHARACTERISTICS:  
;; LENGTH: 653 base pairs  
;; TYPE: nucleic acid  
;; STRANDEDNESS: single  
;; TOPOLOGY: linear

;; MOLECULE TYPE: DNA (genomic)  
;; US-08-455-736-103  
;;  
;; Alignment Scores:  
;; Pred. No.: 7,81e-57 Length: 653  
;; Score: 473.50 Matches: 90  
;; Percent Similarity: 88.33% Conservative: 16  
;; Best Local Similarity: 75.00% Mismatches: 10  
;; Query Match: 78.01% Indels: 4  
;; DB: 2 Gaps: 1  
;;  
;; US-09-786-715-4 (1-118) x US-08-455-736-103 (1-653)  
;;  
QY 3 SerSerGluGluGlyGlnValIleGlyCysHisSerValAspLysTrpProHisValIle 22  
DB 69 TCATCCGAGGAGGACAGTTCGGCTGCCACAGGTTGAGGATGGAAGCAGTACTTC 128  
QY 23 GlnLysGlyValAspSerLysLysLeuValIleAspPheThrAlaSerTrpCysGly 42  
DB 129 AAGAAAGCGCTGAGCTAAGAACTGGTGGTGGCTGCTTACTGCTTATGTCGGG 188  
QY 43 ProCysArpPheIleAlaProIleLeuAlaGluMetAlaLysTrpProHisValIle 62  
DB 189 CTTGCCGTTTATTCGCCCAATCTTCGTCGATTCCTAAGAAAGATGCCCATGTATA 248  
QY 63 PheLeuLysValAspValAspGluLeuLysThrValAla-GluGluPheLysValGlu 82  
DB 249 TTCCTCAAGGTGATGTGATGAGCACTGAAGACTGTTCAACGGGAAATGAGTGGAGGC 308  
QY 82 aMetProThrPheValPheLeuLysGluValGluValGluValGluValGluVal 102  
DB 309 AATGCCAAGCTTTGCTCTTAAGATGGAAGAAGTGACAGAGTTGTGTCGCA 368  
QY 102 GlyGluGluLeuGluAlaThrValGluLysHisGlyAla-----IleThrAla 118  
DB 369 GAAAGAGAGTTCACAGACCATATGAAAGCATGCTGCTCGTCACTGCTACTCT 426  
;;  
RESULT 13  
;; US-08-971-217-103  
;; Sequence 103, Application US/08971217  
;; Patent No. 5942652  
;;  
;; GENERAL INFORMATION:  
;; APPLICANT: Ryals, John A.  
;; APPLICANT: Harms, Christian  
;; APPLICANT: Friedlich, Leslie  
;; APPLICANT: Beck, James  
;; APPLICANT: Uknes, Scott  
;; APPLICANT: Ward, Eric  
;; TITLE OF INVENTION: INDUCIBLE HERBICIDE RESISTANCE  
;; NUMBER OF SEQUENCES: 111  
;; CORRESPONDENCE ADDRESS:  
;; ADDRESSEE: No. 594262artis Corporation  
;; STREET: 3054 Cornwallis Road, P.O. Box 12257  
;; CITY: Research Triangle Park  
;; STATE: NC  
;; COUNTRY: USA  
;; ZIP: 27709  
;;  
;; COMPUTER READABLE FORM:  
;; MEDIUM TYPE: Floppy disk  
;; COMPUTER: IBM PC compatible  
;; OPERATING SYSTEM: PC-DOS/MS-DOS  
;; SOFTWARE: Patentin Release #1.0, Version #1.25  
;; CURRENT APPLICATION DATA:  
;; APPLICATION NUMBER: US/08/971,217  
;; FILING DATE:  
;; CLASSIFICATION: 800  
;; PRIOR APPLICATION DATA:  
;; APPLICATION NUMBER: US 08/457,364  
;; FILING DATE: 31-MAY-1995  
;; PRIOR APPLICATION DATA:  
;; APPLICATION NUMBER: US 08/181,271  
;; FILING DATE: 13-JAN-1994  
;; PRIOR APPLICATION DATA:



```

: FILING DATE: 6-APR-1993
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: US 07/632,441
: FILING DATE: 21-DEC-1990
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: US 07/425,504
: FILING DATE: 20-OCT-1989
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: US 07/848,506
: FILING DATE: 6-MAR-1992
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: US 07/768,122
: FILING DATE: 27-SEP-1991
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: US 07/580,431
: FILING DATE: 7-SEP-1990
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: US 07/368,672
: FILING DATE: 20-JUN-1989
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: US 07/329,018
: FILING DATE: 24-MAR-1989
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: US 07/381,443
: FILING DATE: 18-JUL-1989
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: US 07/353,312
: FILING DATE: 17-MAY-1989
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: US 07/226,303
: FILING DATE: 29-JUL-1988
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: US 08/045,957
: FILING DATE: 12-APR-1993
: ATTORNEY/AGENT INFORMATION:
: NAME: Meigs, J. Timothy
: REGISTRATION NUMBER: 38,241
: REFERENCE/DOCKET NUMBER: S-198250
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: (919)541-8587
: TELEFAX: (919)541-8689
: INFORMATION FOR SEQ ID NO: 103:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 653 base pairs
: TYPE: nucleic acid
: STRANDEDNESS: single
: TOPOLOGY: linear
: MOLECULE TYPE: DNA (genomic)
: US-09-350-600-103

Alignment Scores:
Pred. No.: 7.81e-57 Length: 653
Score: 473.50 Matches: 90
Percent Similarity: 88.33% Conservative: 16
Best Local Similarity: 75.00% Mismatches: 10
Query Match: 78.01% Indels: 4
DB: 3 Gaps: 1

US-09-786-715-4 (1-118) x US-09-350-600-103 (1-653)
Oy 3 SerSerGlugluGlnValIleGlyCysHisserValAspGIUTPrpLysGluGlnPhe 22
Db 69 TCATCCGAGGAGGACAAAGTTCGCTGCCACAAAGTTGAGGAATGAGACGAGTACTTCC 128
Oy 23 GlnLysGlyValAspSerLysLysLeuValValIleAspPheThrAlaSerTrpCysGly 42
Db 129 AAGAAAGCGCTTGAGACTAAGAACTGCTGCTGCTCATTTTACTGCTTCAATGGTGGCGS 188
Oy 43 ProCysArpPheIleAlaProIleLeuAlaGluMetAlaLysLysTrpProHisValIle 62
Db 189 CTTGGCGTTTATTGGCCCAATCTCTGTCGACATCTCTAAGAAAGATGCCCAATGTATA 248
Oy 63 PheLeuLysValAspValAspGluLeuLysThrValAla-GluGluPheLysValGluAl 82
```

```

Db 249 TTCCTCAAGGTTGATGTTGAGAACTGTTTCAGCGGGAATGGAGTGTGAGGC 308
Oy 82 aMetProThrPheValPheLeuLysGluGlyValGluArgLeuValGlyAlaAr 102
Db 309 AATGCCAAGTTTCTTCTTCATTAAGATGAAAGAAAGTGCACAGATGTTGGTGC 368
Oy 102 gLysGluGluLeuGlnAlaThrValGluLysHisGlyAla-----IleThrAla 118
Db 369 GAAGAGAGAGTTGCAGCAGACCAATAGTAAGCATGCTCTCTGCTACTGTACTGCT 426

RESULT 15
US-09-540-014-1
: Sequence 1, Application US/09540014
: Patent No.: 6380372
: GENERAL INFORMATION:
: APPLICANT: Cho, Myeong-Je
: APPLICANT: Del Val, Greg
: APPLICANT: Calileu, Maxime
: APPLICANT: Lemaux, Peggy G.
: APPLICANT: Buchanan, Bob B.
: TITLE OF INVENTION: Barley Gene for Thioredoxin and
: TITLE OF INVENTION: NADP-Thioredoxin Reductase
: FILE REFERENCE: 2001-0701.30
: CURRENT APPLICATION NUMBER: US/09/540,014
: CURRENT FILING DATE: 2000-03-31
: PRIOR APPLICATION NUMBER: US 60/127,198
: PRIOR FILING DATE: 1999-03-31
: PRIOR APPLICATION NUMBER: US 60/169,162
: PRIOR FILING DATE: 1999-12-06
: PRIOR APPLICATION NUMBER: US 60/177,740
: PRIOR FILING DATE: 2000-01-21
: PRIOR APPLICATION NUMBER: US 60/177,739
: PRIOR FILING DATE: 2000-01-21
: NUMBER OF SEQ ID NOS: 51
: SOFTWARE: FastSeq for Windows Version 4.0
: SEQ ID NO 1
: LENGTH: 369
: TYPE: DNA
: ORGANISM: Artificial Sequence
: FEATURE:
: OTHER INFORMATION: barley thioredoxin h cDNA
: US-09-540-014-1

Alignment Scores:
Pred. No.: 1.21e-41 Length: 369
Score: 362.00 Matches: 67
Percent Similarity: 77.19% Conservative: 21
Best Local Similarity: 58.77% Mismatches: 26
Query Match: 59.64% Indels: 0
DB: 4 Gaps: 0

US-09-786-715-4 (1-118) x US-09-540-014-1 (1-369)
Oy 2 AlaSerSerGlugluGlnValIleGlyCysHisserValAspGIUTPrpLysGluGln 21
Db 19 GCGGCGGACAGTGGCGGCGGAGAGTGTCTCGTCCACAGCCGTGAGCAGTGCACATG 78
Oy 22 PheGlnLysGlyValAspSerLysLysLeuValValIleAspPheThrAlaSerTrpCys 41
Db 79 ATCGAGAGAGGCCAACACCCGCAAGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 138
Oy 42 GlyProCysArpPheIleAlaProIleLeuAlaGluMetAlaLysLysTrpProHisVal 61
Db 139 GCACCAATGCGCATCATGCGCTCCAGTTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 198
Oy 62 IlePheLeuLysValAspValAspGluLeuLysThrValAlaGluGluPheLysValGlu 81
Db 199 GTTTTCTCAAGGTCGACGAGGAGTGAAGCCCTTCTGCTGCTGCTGCTGCTGCTGCTG 258
Oy 82 AlaMetProThrPheValPheLeuLysGluGlyValGluArgLeuValGlyAla 101
Db 259 GCCATGCCAACGTTCTCTGCTTCTGATGAAAGAAAGAGACGTCAGAGGAGGTTTGGAG 318
```



QY 102 ArgLysGluLueGlnAlaThrValGluLysHisGlyAla 115  
|||||  
Db 319 ATCAAGAGAGAACTGACCGCCCAAGGTTGGGCTTTCACGGCG 360

Search completed: August 17, 2003, 21:45:13  
Job time : 49.7308 secs



GenCore version 5.1.6  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - nucleic search, using frame\_plus\_p2n model

Run on: August 17, 2003, 19:25:47 ; Search time 137.545 Seconds

(without alignments) 1919.938 Million cell updates/sec

Title: US-09-786-715-4

Perfect score: 607

Sequence: 1 MASSEGGVICHSDENKE.....VGARKEELQATYKGAITTA 118

Scoring table:  
BLOSUM62  
Xgapop 10.0 , Xgapext 0.5  
Ygapop 10.0 , Ygapext 0.5  
Fgapop 6.0 , Fgapext 7.0  
Delop 6.0 , Delext 7.0

Searched: 1504479 seqs, 1118970152 residues

Total number of hits satisfying chosen parameters: 3008958

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Command line parameters: -DEV-xlp  
-MODEL-frame-p2n.model -DEV-xlp  
-O-/cgn2\_1/USPTO\_spool/US09786715/runatc\_11082003\_150515\_6129/app\_query.fasta\_1.1052  
-DB-Published\_Applications\_NA -OFMT=fastap -SUFFIX=rrpb -MINMATCH=0.1  
-LOOPEXT=0 -UNITS=bits -SPART=1 -END=1 -MATRIX=blomsu62  
-FRAN=human40.cdi -LIST=45 -DOCALLIGN=200 -THR SCORE=pcpt -THR MAX=100  
-THR MIN=0 -ALIGN=15 -MODE=LOCAL -OUTFMT=pcpt -NORM=ext -HEAPSIZE=500 -MILEN=0  
-MAXLEN=2000000000 -USER=US09786715\_@CGN\_1.1.560\_@runatc\_11082003\_150515\_6129  
-NCPU=6 -ICPU=3 -NO\_MAP -LARGEDEVY -NEG\_SCORES=0 -WAIT -DSPBLOCK=100  
-LONGLOG -DEV\_TIMEOUT=120 -WARN\_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5  
-XGAPOP=6 -XGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

Published Applications\_NA.\*  
1: /cgn2\_6/ptodata/2/pubpna/US07\_PUBCOMB.seq.\*  
2: /cgn2\_6/ptodata/2/pubpna/PCCT\_NEW\_PUB.seq.\*  
3: /cgn2\_6/ptodata/2/pubpna/US06\_NEW\_PUB.seq.\*  
4: /cgn2\_6/ptodata/2/pubpna/US06\_PUBCOMB.seq.\*  
5: /cgn2\_6/ptodata/2/pubpna/US07\_NEW\_PUB.seq.\*  
6: /cgn2\_6/ptodata/2/pubpna/PCCTUS\_PUBCOMB.seq.\*  
7: /cgn2\_6/ptodata/2/pubpna/US08\_NEW\_PUB.seq.\*  
8: /cgn2\_6/ptodata/2/pubpna/US08\_PUBCOMB.seq.\*  
9: /cgn2\_6/ptodata/2/pubpna/US09A\_PUBCOMB.seq.\*  
10: /cgn2\_6/ptodata/2/pubpna/US09B\_PUBCOMB.seq.\*  
11: /cgn2\_6/ptodata/2/pubpna/US09C\_PUBCOMB.seq.\*  
12: /cgn2\_6/ptodata/2/pubpna/US09D\_NEW\_PUB.seq.\*  
13: /cgn2\_6/ptodata/2/pubpna/US10A\_PUBCOMB.seq.\*  
14: /cgn2\_6/ptodata/2/pubpna/US10B\_PUBCOMB.seq.\*  
15: /cgn2\_6/ptodata/2/pubpna/US10C\_NEW\_PUB.seq.\*  
16: /cgn2\_6/ptodata/2/pubpna/US60\_NEW\_PUB.seq.\*  
17: /cgn2\_6/ptodata/2/pubpna/US60\_PUBCOMB.seq.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

Result Query 8  
No. Score Match Length DB ID Description  
-----  
1 437 72.0 392 10 US-09-878-574-104 Sequence 104, App

2 434 71.5 402 10 US-09-878-574-3723 Sequence 3723, Ap  
3 430 70.8 390 10 US-09-878-574-1462 Sequence 1462, Ap  
4 420 69.2 3888 9 US-09-897-898-10 Sequence 10, Appl  
5 419 69.0 345 10 US-09-938-842A-163 Sequence 163, App  
6 419 69.0 3129 9 US-09-897-898-5 Sequence 5, Appl1  
7 419 69.0 3888 9 US-09-897-898-7 Sequence 7, Appl1  
8 390 64.3 540 12 US-10-349-782-12 Sequence 12, Appl  
9 389 64.1 357 10 US-09-938-842A-1736 Sequence 1736, Ap  
10 389 64.1 357 14 US-10-322-362-1 Sequence 1, Appl1  
11 374.5 61.7 360 10 US-09-938-842A-164 Sequence 164, App  
12 362 59.6 369 12 US-10-194-885-8 Sequence 8, Appl1  
13 362 59.6 369 12 US-10-091-841-1 Sequence 1, Appl1  
14 361 59.5 659 12 US-10-349-782-9 Sequence 9, Appl1  
15 358 59.0 382 12 US-10-091-841-3 Sequence 3, Appl1  
16 358 59.0 393 12 US-10-091-841-5 Sequence 5, Appl1  
17 354 53.4 260 10 US-09-878-574-12941 Sequence 12941, A  
18 324 53.4 270 10 US-09-878-574-13445 Sequence 13445, A  
19 321 52.9 277 10 US-09-878-574-8135 Sequence 8135, Ap  
20 309 50.9 370 10 US-09-878-574-3169 Sequence 3169, Ap  
21 305 50.2 268 10 US-09-878-574-5995 Sequence 5995, Ap  
22 301 49.6 272 10 US-09-878-574-12181 Sequence 12181, A  
23 295 48.6 402 10 US-09-938-842A-225 Sequence 225, App  
24 295 48.6 613 9 US-09-770-149-855 Sequence 855, App  
25 286 47.1 250 9 US-09-923-876-1288 Sequence 1288, Ap  
26 285 47.0 267 10 US-10-349-782-113137 Sequence 13137, A  
27 279 46.0 297 12 US-10-349-782-11 Sequence 11, Appl  
28 279 46.0 456 10 US-09-924-035A-677 Sequence 677, App  
29 256 42.2 390 10 US-09-938-842A-2688 Sequence 2688, Ap  
30 249 41.0 413 14 US-10-102-524-1396 Sequence 1396, Ap  
31 249 41.0 418 10 US-09-796-692-3641 Sequence 3641, Ap  
32 249 41.0 418 14 US-10-040-862-3641 Sequence 3641, Ap  
33 249 41.0 464 10 US-09-998-598-1059 Sequence 1059, Ap  
34 249 41.0 479 11 US-09-535-459-1145 Sequence 1145, Ap  
35 249 41.0 536 14 US-10-066-543-65 Sequence 65, Appl  
36 249 41.0 540 10 US-09-920-300A-1247 Sequence 1247, Ap  
37 249 41.0 557 13 US-10-033-528-1247 Sequence 1247, Ap  
38 249 41.0 557 10 US-09-884-441-88 Sequence 88, Appl  
39 249 41.0 557 11 US-09-907-969-88 Sequence 88, Appl  
40 249 41.0 557 14 US-10-198-053-88 Sequence 88, Appl  
41 249 41.0 568 11 US-09-535-459-1139 Sequence 1139, Ap  
42 249 41.0 570 11 US-09-535-459-1149 Sequence 1149, Ap  
43 249 41.0 594 10 US-09-884-441-87 Sequence 87, Appl  
44 249 41.0 594 11 US-09-907-969-87 Sequence 87, Appl  
45 249 41.0 594 14 US-10-198-053-87 Sequence 87, Appl

#### ALIGNMENTS

RESULT 1  
US-09-878-574-104  
Sequence 104, Application US/09878574  
Patent No. US20020110548A1  
GENERAL INFORMATION:  
APPLICANT: BYRUM, Joseph R.  
APPLICANT: Ia Rosa, Thomas J.  
APPLICANT: Thompson, Michael D.  
TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated with  
TITLE OF INVENTION: Plants  
FILE REFERENCE: 38-21(15401)B  
CURRENT APPLICATION NUMBER: US/09/878,574  
PRIORITY FILING DATE: 2001-12-21  
PRIOR APPLICATION NUMBER: 09/333,535  
PRIORITY FILING DATE: 1999-06-14  
NUMBER OF SEQ ID NOS: 15775  
SEQ ID NO 104  
LENGTH: 392  
TYPE: DNA  
ORGANISM: Glycine max  
OTHER INFORMATION: Clone ID: LIB3028-056-01-B1-B3  
US-09-878-574-104  
Alignment Scores: 1.82e-53 Length: 392  
Pred. No.:









```

: LENGTH: 357
: TYPE: DNA
: ORGANISM: Arabidopsis thaliana
US-10-323-362-1

Alignment Scores:
Pred. No.: 1,25e-46 Length: 357
Score: 389.00 Matches: 70
Percent Similarity: 80.53% Conservative: 21
Best Local Similarity: 61.95% Mismatches: 22
Query Match: 64.09% Indels: 0
DB: 14 Gaps: 0

US-09-786-715-4 (1-118) x US-10-323-362-1 (1-357)

OY 6 GluclglnValIleGlyCysHisSerValaspGluTrpLysGluInpGlnIleGlySgLy 25
Db 10 GAAGGAGAGATTATCGCTTGGCCACACCGTTGAAGATTGGACCGAGAGCTCAAGCCGCC 69
OY 26 ValaspSerLysLsLeuValValIleaspPheThrIleAserTPCySgLyProCySarg 45
Db 70 AACGAATCCAGAAACATGATGTGATGATGACTTCACTCACTGCAATATGTGTGCCCTTGGCGT 129
OY 46 PheIlealProlIleuAlaGluMetalAlaLysLysThrProHisValIlePheLys 65
Db 130 TTCAATTGACCCGCTTTGCTTGACTTGGCCCAAGAACCTCGACAGTAGCTTCTTCAAG 189
OY 66 ValaspValaspGluLeuLysThrValAlaGluGluPheLysValGluAlaMetProThr 85
Db 190 GTTCATGTTGACGAATTGACACATGCTGTGAGAGATTAAAGTTCCAGGCCAATGCCAAGC 249
OY 86 PheValPheLeuLysGluGlyLysGluValGluArgLeuValAlaArgLysGluGlu 105
Db 250 TTATCTTCATGAAGAAGAGAGAGATCAAGAGACTGTGTTGTGCTGCTAAGAGAGAA 309
OY 106 LeuGlnAlaThrValGluLysHisGlyAlaIleThrAla 118
Db 310 ATCAATTGCCAATCTCGAGAGACCAAGACAGACTGTGCT 348

RESULT 11
US-09-938-842A-164
: Sequence 164, Application US/09938842A
: Patent No. US20020160378A1
: GENERAL INFORMATION:
: APPLICANT: Harper, Jeff
: APPLICANT: Kreps, Joel
: APPLICANT: Wang, Xun
: APPLICANT: Zhu, Tong
: TITLE OF INVENTION: STRESS-REGULATED GENES OF PLANTS, TRANSGENIC PLANTS CONTAINING
: TITLE OF INVENTION: SAME, AND METHODS OF USE
: FILE REFERENCE: SCRIPI300-3
: CURRENT APPLICATION NUMBER: US/09/938, 842A
: CURRENT FILING DATE: 2001-08-24
: PRIOR APPLICATION NUMBER: US 60/227, 866
: PRIOR FILING DATE: 2000-08-24
: PRIOR APPLICATION NUMBER: US 60/264, 647
: PRIOR FILING DATE: 2001-01-16
: PRIOR APPLICATION NUMBER: US 60/300, 111
: PRIOR FILING DATE: 2001-06-22
: NUMBER OF SEQ ID NOS: 5379
: SEQ ID NO 164
: LENGTH: 360
: TYPE: DNA
: ORGANISM: Arabidopsis thaliana
US-09-938-842A-164

Alignment Scores:
Pred. No.: 1.53e-44 Length: 360
Score: 374.50 Matches: 72
Percent Similarity: 76.72% Conservative: 17
Best Local Similarity: 62.07% Mismatches: 26
Query Match: 61.70% Indels: 1
DB: 10 Gaps: 1

```

```

US-09-786-715-4 (1-118) x US-09-938-842A-164 (1-360)

OY      3 SerSerIgtugIugIynValIIegIyCySHisSerValAspIuTrpIySgluInPhe 22
Db      4 GCGGCAAGAGGGGTCAAGTGAATGGTTGTTCACACGCAACATGTATGAGCTGTGCAACT 63
OY      23 GlnIySglValAspSerLysLysLeuValValIIeAspPheThrAlaSerTrpCysGly 42
Db      64 GATTAAGCCAAAGATTCACACAGCTGATTTGTGATTGATTTCACCTGCTTCATGTGTCCA 12
OY      43 ProCysArpPheIIeaIaProIIeLeuAlaGluMetAlaLysLys--ThrProHisVal 61
Db      124 CCATGCCGCAGATGATGTCTCCAAATTTTCAACGATTTGGCCAAAGATTCAATGCAACTGCC 18
OY      62 IIePheLeuLysValAspValAspValAspGluLeuLysThrValAlaIaIuGluPheLysValGlu 81
Db      184 ATCTTCTTCAAGAGGGATGTTGATGAACCTTCACAGAGTTCCTTAAGAAGTTGGTGTGGAG 24
OY      82 AlaMetProThrPheValPheLeuLysGluIyLysGluValGluArgLeuValGlyVala 10
Db      244 GCATTTGCCAACCTTTGTGTTCATTATAAAGCCGCGCAAGTTGTGATTAAGCTCGTTGGTGGC 30
OY      102 ArgIySgluIuLeuGlnAlaIaThrValGluLysHisGlyAlaIleThr 117
Db      304 AATAAGAGAGATCTTCAGGCGGAAATATGTGAAGCATCTGCTGTTCACA 351

RESULT 12
US-10-194-885-8
; Sequence 8, Application US/10194885
; Publication No. US20030135878A1
; GENERAL INFORMATION:
; APPLICANT: Wong, J. H.
; APPLICANT: Cho, Myeong-Je
; APPLICANT: Lemaux, Peggy G.
; APPLICANT: Buchanan, Bob
; TITLE OF INVENTION: TRANSGENIC PLANTS WITH ELEVATED
; TITLE OF INVENTION: THIOREDOXIN LEVELS
; FILE REFERENCE: 41627200800
; CURRENT APPLICATION NUMBER: US/10/194,885
; CURRENT FILING DATE: 2002-07-12
; PRIOR APPLICATION NUMBER: 60/307,006
; PRIOR FILING DATE: 2001-07-19
; PRIOR APPLICATION NUMBER: 09/538,864
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: 60/126,736
; PRIOR FILING DATE: 1999-03-29
; NUMBER OF SEQ ID NOS: 55
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 8
; LENGTH: 369
; TYPE: DNA
; ORGANISM: Barley
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)...(369)
US-10-194-885-8

Alignment Scores:
Pred. No.:          9,92e-43          Length:          369
Score: No.:         362.00           Matches:          67
Percent Similarity: 77.19%           Conservative:    21
Best Local Similarity: 58.77%         Mismatches:     26
Query Match:        59.64%           Indels:         0
DB:                12                Gaps:           0

US-09-786-715-4 (1-118) x US-10-194-885-8 (1-369)

OY      2 AlaSerSerIgtugIugIynValIIegIyCySHisSerValAspIuTrpIySgluIn 21
Db      19 GCGGCGGCAATGGCGCGGAGGTGATCTCGGTCCACAGGCTGTGAGCAGTGTGACATGCGAG 78
OY      22 PheGlnIySglValAspSerLysLysLeuValValIIeAspPheThrAlaSerTrpCys 41

```



```
Db 79 ATCGAGGAGGCCAACACCGCCCAAGAGCTGGTGGTATTGACTTCCACTGCATCATGGTGC 138
QY 42 GlyProCysArgPheIleAlaProIleLeuAlaGluMetAlaLysLysThrProHisVal 61
Db 139 GGACCATGCCGATCATGGCTCCAGCTTTTCGCTGATCTCGCCCAAGAGTTCGCCAAATGCT 198
QY 62 IlePheLeuLysValAspGluLeuLysThrValAlaGluGluPheLysValGlu 81
Db 199 GTTTCCTCAAGTGCAGCTGATGAAGTGAAGCCCATGCTGACCAATTCAGTGTGAG 258
QY 82 AlaMetProThrPheValPheLeuLysGluGlyLysGluValGluArgLeuValGlyAla 101
Db 259 GCCATGCCAACCTTCTGTTCAATGAAGGAAGGAGAGCTCAAGGACAGAGGTTCGCGAGCT 318
QY 102 ArgLysGluGluLeuGlnAlaThrValGluLysHisGlyAla 115
Db 319 ATCAAGGAGGAAGTCAACGCCCAAGGTTGGGCTTCACGGCGG 360

RESULT 13
US-10-091-841-1
; Sequence 1, Application US/10091841
; Publication No. US20030150010A1
; GENERAL INFORMATION:
; APPLICANT: Cho, Myeong-Je
; APPLICANT: Del Val, Greg
; APPLICANT: Caillau, Maxime
; APPLICANT: Lemaux, Peggy G.
; APPLICANT: Buchanan, Bob B.
; TITLE OF INVENTION: Barley Gene for Thioredoxin and
; FILE OF INVENTION: NADP-Thioredoxin Reductase
; FILE REFERENCE: 2001-0701.30
; CURRENT APPLICATION NUMBER: US/10/091,841
; PRIOR FILING DATE: 2002-03-05
; PRIOR APPLICATION NUMBER: 09/540,014
; PRIOR FILING DATE: 2000-03-31
; PRIOR APPLICATION NUMBER: US 60/127,198
; PRIOR FILING DATE: 1999-03-31
; PRIOR APPLICATION NUMBER: US 60/169,162
; PRIOR FILING DATE: 1999-12-06
; PRIOR APPLICATION NUMBER: US 60/177,740
; PRIOR FILING DATE: 2000-01-21
; PRIOR APPLICATION NUMBER: US 60/177,739
; PRIOR FILING DATE: 2000-01-21
; NUMBER OF SEQ ID NOS: 51
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 369
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: barley thioredoxin h cDNA
US-10-091-841-1

Alignment Scores:
Pred. No.: 9,92e-43 Length: 369
Score: 362.00 Matches: 67
Percent Similarity: 77.19% Conservativeness: 21
Best Local Similarity: 58.77% Mismatches: 26
Query Match: 59.64% Indels: 0
DB: 12 Gaps: 0

US-09-786-715-4 (1-118) x US-10-091-841-1 (1-369)
QY 2 AlaSerSerGluGluGlyGlnValIleGlyCysHisSerValAspGluTrpLysGluGln 21
Db 19 GCGGCGGAGTGGCGCGGAGGTGATCTCGGTCCACAGCTGGAGCAGTGCACATGCGAG 78
QY 22 PheGlnLysGlyValAspSerLysLysLeuValValIleAspPheThrAlaSerTrpCys 41
Db 79 ATCGAGGAGGCCAACACCGCCCAAGAGCTGGTGGTATTGACTTCCACTGCATCATGGTGC 138
QY 42 GlyProCysArgPheIleAlaProIleLeuAlaGluMetAlaLysLysThrProHisVal 61
```

```
Db 139 GSACCATGCCGATCATGGCTCCAGTTCCTTCGCTGATCTCGCCCAAGAGTTCGCCAAATGCT 198
QY 62 IlePheLeuLysValAspGluLeuLysThrValAlaGluGluPheLysValGlu 81
Db 199 GTTTCCTCAAGTGCAGCTGATGAAGTGAAGCCCATGCTGACCAATTCAGTGTGAG 258
QY 82 AlaMetProThrPheValPheLeuLysGluGlyLysGluValGluArgLeuValGlyAla 101
Db 259 GCCATGCCAACCTTCTGTTCAATGAAGGAAGGAGAGCTCAAGGACAGAGGTTCGCGAGCT 318
QY 102 ArgLysGluGluLeuGlnAlaThrValGluLysHisGlyAla 115
Db 319 ATCAAGGAGGAAGTCAACGCCCAAGGTTGGGCTTCACGGCGG 360

RESULT 14
US-10-349-782-9
; Sequence 9, Application US/10349782
; Publication No. US20030143618A1
; GENERAL INFORMATION:
; APPLICANT: Yves Hatzfield
; APPLICANT: Valerie Marie-No. US20030143618A1lle Frankard
; APPLICANT: Anne-Marie Droual
; TITLE OF INVENTION: Method for easy cloning and selection of chimeric DNA molecule
; FILE REFERENCE: 1187-15
; CURRENT APPLICATION NUMBER: US/10/349,782
; PRIOR FILING DATE: 2003-01-23
; PRIOR APPLICATION NUMBER: EP 02075373.7
; PRIOR FILING DATE: 2002-01-23
; NUMBER OF SEQ ID NOS: 12
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 9
; LENGTH: 659
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Thioredoxin of Oryza sativa in vector pDONR201
US-10-349-782-9

Alignment Scores:
Pred. No.: 3.15e-42 Length: 659
Score: 361.00 Matches: 70
Percent Similarity: 78.45% Conservativeness: 21
Best Local Similarity: 60.34% Mismatches: 25
Query Match: 59.47% Indels: 1
DB: 12 Gaps: 0

US-09-786-715-4 (1-118) x US-10-349-782-9 (1-659)
QY 3 SerSerGluGluGlyGlnValIleGlyCysHisSerValAspGluTrpLysGluGlnPhe 22
Db 133 GCGGCGGAGGAGGAGTCTGATCGCTGCCACACAGGACGAGTTCGACGCCCATGATG 192
QY 23 GlnLysGlyValAspSerLysLysLeuValValIleAspPheThrAlaSerTrpCysGly 42
Db 193 ACCAAGGCCAAGGAGGCGCGGAAAGTGTGATGATTCATATTCACCTTCCTGCTGTGTC 252
QY 43 ProCysArgPheIleAlaProIleLeuAlaGluMetAlaLysLysThrProHisValIle 62
Db 253 CCTTCCCGCTTCATCGCCCCAGTTCGCTGTAATACGCCCAAAAGTTCCTCTGGTGTGTC 312
QY 63 PheLeuLysValAspValAspGluLeuLysThrValAlaGluGluPheLysValGluAla 82
Db 313 TTCTCGAAGTGTGATGTTGATGAGCTGAAGGAGTGTCTGAAAGATACATGTCGAGGCA 372
QY 83 MetProThrPheValPheLeuLysGluGlyLysGluValGluArgLeuValGlyAlaArg 102
Db 373 ATGCGGACCTTCCTATT-ATCAAGGATGCTGCTGAGGTGACAAAGTTCCTTGGCCGAGG 431
QY 103 LysGluGluLeuGlnAlaThrValGluLysHisGlyAlaIleThrAla 118
Db 432 AAGGATGACCTCCAGAACCATCTGTAAGCAGCTCGGTGCGCTGCTGCT 479
```

## RESULT 15

US-10-091-841-3  
; Sequence 3, Application US/10091841  
; Publication No. US20030150010A1  
; GENERAL INFORMATION:  
; APPLICANT: Cho, Myeong-Je  
; APPLICANT: Del Val, Greg  
; APPLICANT: Caillau, Maxime  
; APPLICANT: Lemaux, Peggy G.  
; APPLICANT: Buchanan, Bob B.  
; TITLE OF INVENTION: Barley Gene for Thioresdoxin and  
; TITLE OF INVENTION: NADP-Thioresdoxin Reductase  
; FILE REFERENCE: 2001-0701.30  
; CURRENT APPLICATION NUMBER: US/10/091,841  
; CURRENT FILING DATE: 2002-03-05  
; PRIOR APPLICATION NUMBER: 09/540,014  
; PRIOR FILING DATE: 2000-03-31  
; PRIOR APPLICATION NUMBER: US 60/127,198  
; PRIOR FILING DATE: 1999-03-31  
; PRIOR APPLICATION NUMBER: US 60/169,162  
; PRIOR FILING DATE: 1999-12-06  
; PRIOR APPLICATION NUMBER: US 60/177,740  
; PRIOR FILING DATE: 2000-01-21  
; PRIOR APPLICATION NUMBER: US 60/177,739  
; PRIOR FILING DATE: 2000-01-21  
; NUMBER OF SEQ ID NOS: 51  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 3  
; LENGTH: 382  
; TYPE: DNA  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: wheat thioresdoxin h cDNA  
US-10-091-841-3

Alignment Scores:  
Pred. No.: 3-91e-42 Length: 382  
Score: 358.00 Matches: 67  
Percent Similarity: 77.06% Conservative: 17  
Best Local Similarity: 61.47% Mismatches: 25  
Query Match: 58.98% Indels: 0  
DB: 12 Gaps: 0

US-09-786-715-4 (1-118) x US-10-091-841-3 (1-382)

QY	7	GlyGlnValIleGlyCysHisSerValAspGluTrpLysGluGlnPheGlnLysGlyVal	26
DB	52	GGGAGGTGATCTCCCTCCACAGCCCTGGAGCAGTGCACCATGCAGATCGAGGAGGCCAAC	111
QY	27	AspSerLysLysLeuValValIleAspPheThrAlaSerTrpCysGlyProCysArgPhe	46
DB	112	GCGGCCAAGAAGCTGCTGTTGATTGACTTCACCTGCATCATGGTGGGACCATGCCGCATT	171
QY	47	IleAlaProIleLeuAlaGluMetAlaLysLysThrProHisValIlePheLeuLysVal	66
DB	172	ATGGCTCCAAATTTTCGCTGATTCGCAAGAAGTTCACAGCTGCTGTTTCTCTCAAGGTC	231
QY	67	AspValAspGluLeuLysThrValAlaGluGluPheLysValGluAlaMetProThrPhe	86
DB	232	GAGGTGATGAACTGAAGCCCATTCGTGAGCAATTCAGGTGGAGGCCATGCCAACCCTTC	291
QY	87	ValPheLeuLysGluGlyLysGluValGluArgLeuValIleAlaArgLysGluGluLeu	106
DB	292	CTGTTTCATGAAGAAGAGATGTCAGAGACAGGGTTGTCGAGCTATCAAGGAGGAACG	351
QY	107	GlnAlaThrValGluLysHisGlyAla	115
DB	352	ACGACCAAGGTTGGGCTACACGCGCC	378

Search completed: August 17, 2003, 23:43:49  
Job time : 140.545 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - nucleic search, using frame\_plus.p2n model

Run on: August 17, 2003, 19:20:27 ; Search time 1595.81 Seconds  
(without alignments)  
1797.162 Million cell updates/sec

Title: US-09-786-715-4  
Perfect score: 607  
Sequence: 1 MASSEGGVICHSHYDEWKE.....YGARKEELQATVEKGAITA 118

Scoring table: BLOSUM62  
Xgapop 10.0, Xgapext 0.5  
Ygapop 10.0, Ygapext 0.5  
Fgapop 6.0, Fgapext 7.0  
Delop 6.0, Delext 7.0

Searched: 22781392 seqs, 12152238056 residues  
Total number of hits satisfying chosen parameters: 45562784

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 08  
Maximum Match 100%  
Listing first 45 summaries

Command line parameters:  
-Q=/cgn2.1/USPTO\_SPOOL/US09786715/runat\_11082003\_150514\_6048/app\_query.fasta\_1.1052  
-DB=EST -QFMT=fastap -SUFFIX=1st -MINMATCH=0.1 -LOOPEXT=0 -LOOPEXT=0  
-UNITS=bits -START=1 -END=1 -MATRIX=biosum62 -TRANS=human40.cdi -LIST=45  
-DOCALIGN=200 -THR.SCORE=Pct -THR.MAX=100 -THR.MIN=0 -ALIGN=15 -MODE=LOCAL  
-OUTFMT=tbl -NORM=ext -HEADSIZE=500 -MINLEN=0 -MAXLEN=2000000000  
-USER=US09786715\_66CN\_1\_1\_7834\_@runat\_11082003\_150514\_6048 -NCPU=6 -ICPU=3  
-NO\_MAP -LARGEOBERT -NEG.SCOES=0 -WAIT -DSPBLOCK=100 -LONGLOG  
-DEV.TIMEOUT=120 -WARN.TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6  
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEX=7

Database :  
EST:\*  
1: em\_estba:\*  
2: em\_estlum:\*  
3: em\_estlin:\*  
4: em\_estlmu:\*  
5: em\_estrov:\*  
6: em\_estpl:\*  
7: em\_estro:\*  
8: em\_hic:\*  
9: gb\_est1:\*  
10: gb\_est2:\*  
11: gb\_hic:\*  
12: gb\_est3:\*  
13: gb\_est4:\*  
14: gb\_est5:\*  
15: em\_estfun:\*  
16: em\_estom:\*  
17: em\_gss\_hum:\*  
18: em\_gss\_inv:\*  
19: em\_gss\_pln:\*  
20: em\_gss\_vit:\*  
21: em\_gss\_fun:\*  
22: em\_gss\_mam:\*  
23: em\_gss\_mus:\*  
24: em\_gss\_pro:\*  
25: em\_gss\_rtd:\*  
26: em\_gss\_phg:\*  
27: em\_gss\_vr1:\*  
28: gb\_gss1:\*

29: gb\_gss2:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	527	86.8	441	12	BM173053
2	512	84.3	433	12	BM173035
3	492	81.1	452	14	CA524170
4	490	80.7	455	9	A1773303
5	480	80.7	477	9	AM041395
6	480	80.7	505	9	AM094525
7	480	80.7	541	9	A1776110
8	480	80.7	546	9	AM038924
9	480	80.7	547	9	AM038141
10	480	80.7	552	9	AM625331
11	480	80.7	564	9	A1775760
12	480	80.7	567	9	AM037831
13	480	80.7	570	12	BI208736
14	480	80.7	573	9	A1778940
15	480	80.7	573	9	A1781827
16	480	80.7	581	9	AM096472
17	480	80.7	594	9	AM650009
18	480	80.7	615	9	AM621673
19	480	80.7	617	9	A1780240
20	480	80.7	619	12	BI210657
21	480	80.7	622	12	BI206688
22	480	80.7	623	9	AM625957
23	480	80.7	694	9	AM040001
24	480	80.7	694	9	AM040011
25	480	80.7	731	12	BI203833
26	489	80.6	455	10	BF050482
27	485	79.9	565	9	AM040195
28	485	79.9	565	9	AM040201
29	485	79.9	624	13	BQ507669
30	485	79.9	628	10	BG598337
31	485	79.9	648	13	BQ507668
32	482	79.4	380	9	AM038564
33	481	79.2	482	9	AM626018
34	477	78.6	574	12	BM063442
35	477	78.6	575	10	BG130196
36	477	78.6	637	12	BM065241
37	476	78.4	494	12	BM878771
38	476	78.4	579	13	BU692317
39	470	77.4	564	13	BU692458
40	470	77.4	590	12	BG890096
41	466	76.8	508	9	AM625037
42	459	75.6	548	14	CA896876
43	459	75.6	555	14	CA901892
44	459	75.6	601	14	CA896877
45	458	75.5	464	9	AL386021

## ALIGNMENTS

RESULT 1  
BM173053  
LOCUS  
DEFINITION  
900407 Avicennia marina leaf cDNA library Avicennia marina cDNA  
clone Am900407 5' similar to thioredoxin (X58527) of Nicotiana  
tabacum, mRNA sequence.  
ACCESSION  
BM173053  
VERSION  
BM173053.1 GI:17312616  
KEYWORDS  
EST.  
SOURCE  
Avicennia marina  
ORGANISM  
Avicennia marina  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;



Pred. No.: 4.2e-58 Length: 433  
 Score: 512.00 Matches: 96  
 Percent Similarity: 94.74% Conservative: 12  
 Best Local Similarity: 84.21% Mismatches: 5  
 Query Match: 84.35% Indels: 1  
 DB: 12 Gaps: 0

US-09-786-715-4 (1-118) x BM173035 (1-433)

OY 1 MetAlaSerSerGluGluGlyGlnValIleGly-CysHisSerValAspGluTrpLysGln 20  
 DB 71 ATGGCTCTCGTGGAGAGGCTCAAGTGAATGCTGCTCCAGTCCGAGAGGAGGA 130  
 OY 20 uGlnPheGlnValAspSerLysLysValValIleAspPheThrAlaSerTr 40  
 DB 131 GCATTTCAGAAAGGGGCTGAGCTAAAGAGGAGTGAAGTATTTACGGCTTCG 190  
 OY 40 PCysGlyProCysArgPheIleAlaProIleLeuAlaGluMetAlaLysLysThrProH 60  
 DB 191 GTCGGAGCTTCGCTTCATTTGCTCCATTTTGGCTGAGATCCGCAAGAGTCCGCTCA 250  
 OY 60 SVAlIlePheLeuLysValAspValAspGluLeuLysThrValAlaGluGluPheLysVa 80  
 DB 251 TGTGTATTCCTGCAAGGTGATGATGATGATGATGATGATGATGATGATGATGAT 310  
 OY 80 LgluAlaMetProThrPheValPheLeuLysGluGlyLysGluValGluArgLeuValG 100  
 DB 311 TGAAGCTATGCTTACCTTCTGCTCTCAAGAGAGGAAAGAAATGACAGGCTTGTGG 370  
 OY 100 yAlaArgLysGluGluLeuGlnAlaThrValGluLysHis 113  
 DB 371 TGCACAGAGAGAGATCTCTGCGCCAGATCACCAACAT 410

RESULT 3  
 CA524170 452 bp mRNA linear EST 15-NOV-2002  
 LOCUS KS1203303 KS12 Capsicum annuum cDNA, mRNA sequence.  
 DEFINITION CA524170  
 ACCESSION CA524170  
 VERSION CA524170.1 GI:25038242  
 KEYWORDS EST.  
 SOURCE Capsicum annuum  
 ORGANISM Capsicum annuum  
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
 asterids; lamids; Solanales; Solanales; Solanales; Solanales;  
 1 (bases 1 to 452)  
 Lee,S., Kim,S.-Y., Chung,Y.-H., Shin,H.-J., Goh,S.-H., Pai,H.-S.,  
 Hur,C.-G. and Choi,D.  
 Generation of Expressed Sequence Tags from Hot Pepper (Capsicum  
 annuum L.) and Sequence Analysis in Relation to Hypersensitive  
 Response Against Pathogen  
 Unpublished

## JOURNAL

CONTACT: Doll Choi  
 Genome Research Center and National Center for Genome Information  
 Korea Research Institute of Bioscience and Biotechnology  
 P.O. Box 115, Yuseong, Taejeon, 305-600, Republic of Korea  
 Tel: 82-42-860-4340  
 Fax: 82-42-860-4309  
 Email: doll@mail.krrib.re.kr  
 Plate: 033 row: H column: 03.  
 Location/Qualifiers  
 1.452  
 /organism="Capsicum annuum"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:4072"  
 /clone\_lib="KS12"

FEATURES  
 source

BASE COUNT 116 a 75 c 121 g 140 t

Alignment Scores:  
 Pred. No.: 2.09e-55 Length: 452  
 Score: 492.00 Matches: 90  
 Percent Similarity: 92.79% Conservative: 13

Best Local Similarity: 81.08% Mismatches: 8  
 Query Match: 81.05% Indels: 0  
 DB: 14 Gaps: 0

US-09-786-715-4 (1-118) x CA524170 (1-452)

OY 3 SerSerGluGluGlyGlnValIleGlyCysHisSerValAspGluTrpLysGlnPhe 22  
 DB 66 TCATCTGAAGAGAGACAACTTTTGGCTTGCACAGGTTGAGAGTGAATGACCTTC 125  
 OY 23 GlnLysGlyValAspSerLysLysLeuValValIleAspPheThrAlaSerTrpCysGly 42  
 DB 126 AAGAGAGGTGTCAGACTAAATATGTTGGTGTGATTTTACCCATCTCGTCCGCT 185  
 OY 43 ProCysArgPheIleAlaProIleLeuAlaGluMetAlaLysLysThrProHisValIle 62  
 DB 186 CCTTGGCCGTTTATTTCCCAATTTGCTGACATGCTTAAGAGATGCCCCATGTGCANA 245  
 OY 63 PheLeuLysValAspValAspGluLeuLysThrValAlaGluGluPheLysValGluAla 82  
 DB 246 TTCCTCAAGCTTGATTTGATGACTTAAGACTGTTGCAGAGCAATGATGTCAGACT 305  
 OY 83 MetProThrPheValPheLeuLysGluGlyLysGluValGluArgLeuValGluAlaArg 102  
 DB 306 ATGCCAACATTTGCTCTCTTAAGATGCAAGAGTGAATGATGATGATGATGATGATG 365  
 OY 103 LysGluGluLeuGlnAlaThrValGluLysHis 113  
 DB 366 AAGAGAGCTTGCAGCAGCAGCATCTTAAGCAT 398

RESULT 4  
 A1773303 455 bp mRNA linear EST 18-MAY-2001  
 LOCUS EST254403 tomato resistant, Cornell Lycopersicon esculentum cDNA  
 DEFINITION A1773303  
 ACCESSION A1773303  
 VERSION A1773303.1 GI:5271344  
 KEYWORDS EST.  
 SOURCE Lycopersicon esculentum (tomato)  
 ORGANISM Lycopersicon esculentum  
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
 asterids; lamids; Solanales; Solanales; Solanales; Solanales;  
 1 (bases 1 to 455)  
 D'Ascenzo,M., He,X., Lyman,J., Matern,A.L., Vision,T., Holt,I.E.,  
 Liang,F., Upton,J., Ronning,C.M., Craven,M.B., Fujii,C.Y., Bowman  
 C.L., Nierman,W., Fraser,C.M., Venter,J.C., Tanksley,S.D.,  
 Giovannoni,J.J. and Martin,G.B.  
 Generation of ESTs from Pseudomonas resistant tomato  
 Unpublished  
 Other\_ESTS: TCI792  
 CONTACT: CUGI  
 CLEMSON UNIVERSITY GENOMICS INSTITUTE  
 CLEMSON UNIVERSITY  
 100 JORDAN HALL, CLEMSON, SC 29634, USA  
 Email: <http://www.genome.clemson.edu/orders/index.html>  
 5 prime sequence.  
 Location/Qualifiers  
 1.455  
 /organism="Lycopersicon esculentum"  
 /mol\_type="mRNA"  
 /cullivar="RI1-12 (35S::Pto in Rio Grande x Money Maker)"  
 /db\_xref="taxon:4081"  
 /clone\_lib="CLERE3"  
 /tissue\_type="leaf"  
 /dev\_stage="4-week old"  
 /lab\_host="SOLR"  
 /clone\_lib="tomato resistant, Cornell"  
 /note="Vector: pBluescript SK(-), Site\_1: EcoRI, Site\_2:  
 XhoI; cDNA - Tomato Pseudomonas Resistant EST Library.  
 Directionally cloned cDNAs inserted into pBluescript SK(-  
 ) at 5' end with EcoRI and 3' end with XhoI site."

## JOURNAL

CONTACT: CUGI  
 CLEMSON UNIVERSITY GENOMICS INSTITUTE  
 CLEMSON UNIVERSITY  
 100 JORDAN HALL, CLEMSON, SC 29634, USA  
 Email: <http://www.genome.clemson.edu/orders/index.html>  
 5 prime sequence.  
 Location/Qualifiers  
 1.455  
 /organism="Lycopersicon esculentum"  
 /mol\_type="mRNA"  
 /cullivar="RI1-12 (35S::Pto in Rio Grande x Money Maker)"  
 /db\_xref="taxon:4081"  
 /clone\_lib="CLERE3"  
 /tissue\_type="leaf"  
 /dev\_stage="4-week old"  
 /lab\_host="SOLR"  
 /clone\_lib="tomato resistant, Cornell"  
 /note="Vector: pBluescript SK(-), Site\_1: EcoRI, Site\_2:  
 XhoI; cDNA - Tomato Pseudomonas Resistant EST Library.  
 Directionally cloned cDNAs inserted into pBluescript SK(-  
 ) at 5' end with EcoRI and 3' end with XhoI site."

FEATURES  
 source

BASE COUNT 114 a 78 c 125 g 138 t



```

/cultivar="Rio Grande Ptor"
/db_xref="taxon:4081"
/clone="CLET29E19"
/tissue_type="leaf"
/dev_stage="4-6 week old plants"
/lab_host="X11-Blue MRP"
/clone_lib="tomato mixed elicitor, BTI"
/notes="Vector: pBluescript SK(-); Site_1: EcoRI; Site_2:
XhoI; CLET - Inoculated with a variety of disease response
elicitors. Plants exposed to 2,6 dichloroisonicotinic
acid, BTH, jasmonic acid, ethylene, fenthion, ETX,
okadaic acid, or systemin prior to tissue harvest. EcoRI
site was destroyed during cloning."

```

# BASE COUNT

123 a 82 c 136 g 164 t

## Alignment Scores:

Pred. No.: 4,49e-55 Length: 505  
Score: 490.00 Matches: 91  
Percent Similarity: 89.66% Conservative: 13  
Best Local Similarity: 78.458 Mismatches: 12  
Query Match: 80.72% Indels: 0  
DB: 9 Gaps: 0

US-09-786-715-4 (1-118) x AM094525 (1-505)

```

OY 3 SerSerGluGluGlyGlnValIleGlyCysSHSSerValAspGluTrpLysGluGlnPhe 22
    |||||||
DB 30 TCATCTGAGGAGGAGACAAAGTATCGGCTCCACAAAGTTGAGAGATGTCAGAGCTC 89
    |||||||
OY 23 GlnLysGlyValAspSerLysLysLeuValIleAspPheThrLaseTrpCysGly 42
    |||||||
DB 90 CAGAGAGGTGTGAGAGCCAAAACCTGCTGGTGATTTACTCTCTCTCTCTCTCTCTCT 149
    |||||||
OY 43 ProCysArgPheIleAlaProIleLeuAlaGluMetAlaLysLysThrProHisValIle 62
    |||||||
DB 150 CCTTCCGCTTTTATTTATTCCTCCCAATTCCTGACATTCCTAAGAATGCCCCATGTTTANG 209
    |||||||
OY 63 PheLeuLysValAspValAspGluLeuLysThrValAlaGluGluPheLysValGluAla 82
    |||||||
DB 210 TTCCTCAAGGTGATGTTATGATGAACGAAAGTTGTCAGAGAAATGGAATGTGGAGGCA 269
    |||||||
OY 83 MetProThrPheValPheLeuLysGluGlyLysGluValAlaGluGluPheLysValGluAla 102
    |||||||
DB 270 ATGCCAACCTTTGTCTCTCATTTAAAGAGGTTAAAGAGTGAATGATGATGATGATGATG 329
    |||||||
OY 103 LysGluGluLeuGlnAlaThrValGluLysSHSgLyAlaIleThrAla 118
    |||||||
DB 330 AAAGACGATGCTTCAGACCATAGAGACATGCTGCTCTCTCTCTCTCTCTCTCTCTCTCT 377
    |||||||

```

## RESULT 7

AI776110 541 bp mRNA linear EST 18-MAY-2001

LOCUS EST57198 tomato resistant, Cornell Lycopersicon esculentum CDNA  
DEFINITION clone CLE1702, mRNA sequence.

ACCESSION AI776110  
VERSION AI776110.1 GI:5274139

KEYWORDS EST.  
SOURCE Lycopersicon esculentum (tomato)

ORGANISM Lycopersicon esculentum  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
asterids; lamids; Solanales; Solanaceae; Solanum; Lycopersicon.

REFERENCE 1 (bases 1 to 541)  
D'Ascenzo,M., He,X., Lyman,J., Matern,A.L., Vision,T., Holt,I.E.,  
Liang,F., Upton,J., Ronning,C.M., Craven,M.B., Fujii,C.Y., Bowman,  
C.L., Nierman,M., Fraser,C.M., Venter,J.C., Tanksley,S.D.,  
Giovannoni,J.J. and Martin,G.B.  
Generation of ESTs from Pseudomonas resistant tomato

AUTHORS

REFERENCE

TITLE

JOURNAL

COMMENT  
Contact: CGI  
Clemson University Genomics Institute  
Clemson University

## FEATURES

source  
100 Jordan Hall, Clemson, SC 29634, USA  
Email: <http://www.genome.clemson.edu/orders/index.html>  
5 prime sequence.  
Location/Qualifiers  
1. 541  
/organism="Lycopersicon esculentum"  
/mol\_type="mRNA"  
/cultivar="R11-12 (35S::Pro in Rio Grande x Money Maker)"  
/db\_xref="taxon:4081"  
/clone="CLET1702"  
/tissue\_type="leaf"  
/dev\_stage="4-week old"  
/lab\_host="SOLR"  
/clone\_lib="tomato resistant, Cornell"  
/notes="Vector: pBluescript SK(-); Site\_1: EcoRI; Site\_2:  
XhoI; CLET - Tomato Pseudomonas Resistant EST Library.  
directionally cloned cDNAs inserted into pBluescript SK(-)  
) at 5' end with EcoRI and 3' end with XhoI site."

## BASE COUNT

129 a 84 c 142 g 186 t

## Alignment Scores:

Pred. No.: 4,92e-55 Length: 541  
Score: 490.00 Matches: 91  
Percent Similarity: 89.66% Conservative: 13  
Best Local Similarity: 78.458 Mismatches: 12  
Query Match: 80.72% Indels: 0  
DB: 9 Gaps: 0

US-09-786-715-4 (1-118) x AI776110 (1-541)

```

OY 3 SerSerGluGluGlyGlnValIleGlyCysSHSSerValAspGluTrpLysGluGlnPhe 22
    |||||||
DB 75 TCATCTGAGGAGGAGACAAAGTATCGGCTCCACAAAGTTGAGAGATGTCAGAGCTC 134
    |||||||
OY 23 GlnLysGlyValAspSerLysLysLeuValIleAspPheThrLaseTrpCysGly 42
    |||||||
DB 135 CAGAGAGGTGTGAGAGCCAAAACCTGCTGGTGATTTACTCTCTCTCTCTCTCTCTCT 194
    |||||||
OY 43 ProCysArgPheIleAlaProIleLeuAlaGluMetAlaLysLysThrProHisValIle 62
    |||||||
DB 195 CCTTCCGCTTTTATTTATTCCTCCCAATTCCTGACATTCCTAAGAATGCCCCATGTTTANG 254
    |||||||
OY 63 PheLeuLysValAspValAspGluLeuLysThrValAlaGluGluPheLysValGluAla 82
    |||||||
DB 255 TTCCTCAAGGTGATGTTATGATGAACGAAAGTTGTCAGAGAAATGGAATGTGGAGGCA 314
    |||||||
OY 83 MetProThrPheValPheLeuLysGluGlyLysGluValAlaGluGluPheLysValGluAla 102
    |||||||
DB 315 ATGCCAACCTTTGTCTCTCATTTAAAGAGGTTAAAGAGTGAATGATGATGATGATGATG 374
    |||||||
OY 103 LysGluGluLeuGlnAlaThrValGluLysSHSgLyAlaIleThrAla 118
    |||||||
DB 375 AAAGACGATGCTTCAGACCATAGAGACATGCTGCTCTCTCTCTCTCTCTCTCTCTCTCT 422
    |||||||

```

## RESULT 8

AM038924 546 bp mRNA linear EST 18-MAY-2001

LOCUS EST280880 tomato mixed elicitor, BTI Lycopersicon esculentum CDNA  
DEFINITION clone CLET10G5, mRNA sequence.

ACCESSION AM038924  
VERSION AM038924.1 GI:5897678

KEYWORDS EST.  
SOURCE Lycopersicon esculentum (tomato)

ORGANISM Lycopersicon esculentum  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
asterids; lamids; Solanales; Solanaceae; Solanum; Lycopersicon.

REFERENCE 1 (bases 1 to 546)  
D'Ascenzo,M., He,X., Lyman,J., Holt,I.E., Liang,F., Upton,J.,  
Ronning,C.M., Craven,M.B., Fujii,C.Y., Bowman,C.L., Nierman,M.,  
Fraser,C.M., Venter,J.C., Martin,G.B., Tanksley,S.D. and Giovannoni,  
J.J.

AUTHORS

REFERENCE

TITLE

JOURNAL

COMMENT  
Contact: CGI  
Clemson University Genomics Institute  
Clemson University

TITLE Generation of ESTs from tomato leaf tissue  
JOURNAL Unpublished  
COMMENT Contact: CUGI  
Clemson University Genomics Institute  
Clemson University  
100 Jordan Hall, Clemson, SC 29634, USA  
Email: <http://www.genome.clemson.edu/orders/index.html>  
5 prime sequence.

FEATURES  
source Location/Qualifiers  
1..546  
/organism="Lycopersicon esculentum"  
/mol\_type="mRNA"  
/cultivar="Rio Grande ProR"  
/db\_xref="taxon:4081"  
/clone="cLET10G5"  
/tissue\_type="leaf"  
/dev\_stage="4-6 week old plants"  
/lab\_host="X11-Blue MRF"  
/clone\_lib="tomato mixed elicitor, BRT"  
/note="Vector: pBluescript SK(-); Site\_1: EcoRI; Site\_2: XhoI; cLET - Inoculated with a variety of disease response elicitors. Plants exposed to 2,6 dichloroisocortinic acid, BTH, jasmonic acid, ethylene, fenthion, EIX, okadaic acid, or systemin prior to tissue harvest. EcoRI site was destroyed during cloning."

BASE COUNT 129 a 84 c 145 g 188 t

ORIGIN

Alignment Scores:  
Pred. No.: 4.98e-55 Length: 546  
Score: 490.00 Matches: 91  
Percent Similarity: 89.66% Conservative: 13  
Best Local Similarity: 78.45% Mismatches: 12  
Query Match: 80.72% Indels: 0  
DB: 9 Gaps: 0

US-09-786-715-4 (1-118) x AM038924 (1-546)

QY 3 SerSerGluGluGlnGlnValIleGlyCysHisSerValAspGluTrpLysGluGlnPhe 22  
|||||  
Db 77 TCATCTGAGGAGACACAGTATGGCTGCCACAGGTGAGAGTGGAGAGTGCACCTC 136  
23 GlnLysGlyValAspSerLysLysLeuValIleAspPheThrAlaSerTrpCysGly 42  
|||||  
Db 137 CAGAGGGGTGGGAGACCAAAACTGGTGGTGGATTTACTGCTTCCTGGCGGT 196  
43 ProCysArgPheIleAlaProIleLeuAlaGluMetAlaLysLysTrpProHisValIle 62  
|||||  
Db 197 CCTTGGCGTTTATTTGGCCCAATTCCTGTCGACATTCCTAAGAGATGCCCATGTTATG 256  
63 PheLeuLysValAspValAspGluLeuLysThrValAlaGluGluPheLysValGluAla 82  
|||||  
Db 257 TTCTCTCAAGGTGATGTGATGACAGTGAAGAACTTCAGAGAGTGAATGATGCGAGCA 316  
83 MetProThrPheValPheLeuLysGluGlyLysGluValGluArgLeuValGlyAlaArg 102  
|||||  
Db 317 ATGCAACCTTTGCTTCATTAAAGAGGTAAGAAAGTGAAGTGAAGTGGTGGTGGCAAT 376  
103 LysGluGluLeuGlnAlaThrValGluLysHisGlyAlaIleThrAla 118  
|||||  
Db 377 AAAGACGATTTGCTTCAGACCATAGAGACATGCTGCTCTGCT 424

RESULT 9  
AM038141 547 bp mRNA linear EST 18-MAY-2001  
LOCUS EST279798 tomato mixed elicitor, BRT Lycopersicon esculentum CDNA  
DEFINITION clone cLETIN21, mRNA sequence.  
ACCESSION AM038141  
VERSION AM038141.1 GI:5896895  
KEYWORDS EST.  
SOURCE Lycopersicon esculentum (tomato)  
ORGANISM Lycopersicon esculentum  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;

TITLE Generation of ESTs from tomato leaf tissue  
JOURNAL Unpublished  
COMMENT Contact: CUGI  
Clemson University Genomics Institute  
Clemson University  
100 Jordan Hall, Clemson, SC 29634, USA  
Email: <http://www.genome.clemson.edu/orders/index.html>  
5 prime sequence.

FEATURES  
source Location/Qualifiers  
1..547  
/organism="Lycopersicon esculentum"  
/mol\_type="mRNA"  
/cultivar="Rio Grande ProR"  
/db\_xref="taxon:4081"  
/clone="cLETIN21"  
/tissue\_type="leaf"  
/dev\_stage="4-6 week old plants"  
/lab\_host="X11-Blue MRF"  
/clone\_lib="tomato mixed elicitor, BRT"  
/note="Vector: pBluescript SK(-); Site\_1: EcoRI; Site\_2: XhoI; cLET - Inoculated with a variety of disease response elicitors. Plants exposed to 2,6 dichloroisocortinic acid, BTH, jasmonic acid, ethylene, fenthion, EIX, okadaic acid, or systemin prior to tissue harvest. EcoRI site was destroyed during cloning."

BASE COUNT 129 a 84 c 146 g 188 t

ORIGIN

Alignment Scores:  
Pred. No.: 4.99e-55 Length: 547  
Score: 490.00 Matches: 91  
Percent Similarity: 89.66% Conservative: 13  
Best Local Similarity: 78.45% Mismatches: 12  
Query Match: 80.72% Indels: 0  
DB: 9 Gaps: 0

US-09-786-715-4 (1-118) x AM038141 (1-547)

QY 3 SerSerGluGluGlnGlnValIleGlyCysHisSerValAspGluTrpLysGluGlnPhe 22  
|||||  
Db 77 TCATCTGAGGAGACACAGTATGGCTGCCACAGGTGAGAGTGGAGAGTGCACCTC 136  
23 GlnLysGlyValAspSerLysLysLeuValIleAspPheThrAlaSerTrpCysGly 42  
|||||  
Db 137 CAGAGGGGTGGGAGACCAAAACTGGTGGTGGATTTACTGCTTCCTGGCGGT 196  
43 ProCysArgPheIleAlaProIleLeuAlaGluMetAlaLysLysTrpProHisValIle 62  
|||||  
Db 197 CCTTGGCGTTTATTTGGCCCAATTCCTGTCGACATTCCTAAGAGATGCCCATGTTATG 256  
63 PheLeuLysValAspValAspGluLeuLysThrValAlaGluGluPheLysValGluAla 82  
|||||  
Db 257 TTCTCTCAAGGTGATGTGATGACAGTGAAGAACTTCAGAGAGTGAATGATGCGAGCA 316  
83 MetProThrPheValPheLeuLysGluGlyLysGluValGluArgLeuValGlyAlaArg 102  
|||||  
Db 317 ATGCAACCTTTGCTTCATTAAAGAGGTAAGAAAGTGAAGTGAAGTGGTGGTGGCAAT 376  
103 LysGluGluLeuGlnAlaThrValGluLysHisGlyAlaIleThrAla 118  
|||||  
Db 377 AAAGACGATTTGCTTCAGACCATAGAGACATGCTGCTCTGCT 424

RESULT 10  
AM625331 552 bp mRNA linear EST 18-MAY-2001  
LOCUS EST319154 tomato radicle, 5 d post-imbibition, Cornell University



ACCESSION	AW625331	LYCopersicon esculentum cDNA clone cLEZ12117 5', mRNA sequence.
VERSION	AW625331.1	GI:7338274
KEYWORDS	EST.	
SOURCE	LYCopersicon esculentum (tomato)	
ORGANISM	LYCopersicon esculentum	
REFERENCE	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; asterids; lamiales; Solanales; Solanaceae; Solanum; Lycopersicon. 1 (bases 1 to 352)	
AUTHORS	van der Hoeven,R.S., Garvin,D., Matern,A.L., Holt,I.E., Liang,F., Hansen,T., Craven,M.B., Bowman,C.L., Romling,C.M., Nierman,W., Fraser,C.M., Martin,G.B., Giovannoni,J.J. and Tanksley,S.D.	
TITLE	Generation of ESTs from tomato radicle tissue (etiolated)	
JOURNAL	Unpublished	
COMMENT	Contact: CUGI Clemson University Genomics Institute Clemson University 100 Jordan Hall, Clemson, SC 29634, USA Email: <a href="http://www.genome.clemson.edu/orders/index.html">http://www.genome.clemson.edu/orders/index.html</a> 5 prime sequence.	
FEATURES	Location/Qualifiers	
source	1..552	
	/organism="Lycopersicon esculentum"	
	/mol_type="mRNA"	
	/cultivar="TA496"	
	/db_xref="taxon:4081"	
	/clone="cLEZ12117"	
	/tissue_type="radicle"	
	/dev_stage="seedlings 5 days post-imbibition"	
	/clone.lib="tomato radicle, 5 d post-imbibition, Cornell University"	
	/note="Vector: pBluescript SK(-); Site_1: EcoRI; Site_2: XhoI; supplier: Tanksley; tissue supplied by Dave Garvin (USDA-ARS, Ithaca, NY 14850)."	
BASE COUNT	131 a 86 c 146 g 189 t	
ORIGIN		
Alignment Scores:		
Pred. NO.:	5.05e-55	Length: 552
Score:	490.00	Matches: 91
Percent Similarity:	89.66%	Conservative: 13
Best local Similarity:	78.45%	Mismatches: 12
Query Match:	80.72%	Indels: 0
DB:	9	Gaps: 0
US-09-786-715-4 (1-118) x AW625331 (1-552)		
QY	3 SerSerGIUGlUGlGlnValIleGIYCySHISerValASPGLUtrPlySGlUGlnPhe	22
Db		
78	TCATCTGGAGGAAGGCAAGTGAATCGGCTGCCCAAGAGTGAAGAGTGAAGGTCAGCGTC	137
QY	23 GlnYSGlYValASPserLysLysLeuValIleaspPheThrIasertPrpYSGly	42
Db		
138	CAGAAAGGTGTGGAAACCAAAAACCTGGTGTGGTGGATTTACTGCTTCGTGGTCGGT	197
QY	43 ProCYsArpHeIleAlaProIleLeuAlaGluMetAlaLysLysThrProHisValIle	62
Db		
198	CCTTGCCCTTTAATGCCCAATTCCTTGTCACATTCGTAAGAAGATGCCCATGTTAAG	257
QY	63 PheLeuLysValASPValASPGLULeuLysThrValAlaGluGluPheLysValGluAla	82
Db		
258	TTCCCAAGGTTGATGTGATGACGTGAAGAAAGTTCGACAGAGAAATGATGTGACAGGCA	317
QY	83 MetProThrPheValPheLeuLysGIUGlYLSGIUGlValGluArgLeuValGlyAlaArg	102
Db		
318	ATGCCAACTTTGTGCTTCAITTAAGAGGGTAAGAGTGCATPAGGGTGTGTGGTCCAAAT	377
QY	103 LysGIUGlUGlAlaThrValAlaLysSHISGIYAlaIleThrAla	118
Db		
378	AAAGAGGATTTGCTTCACACCATTAAGAAAGATGTGTGCTGCTGCTGCT	425
RESULT	11	

LOCUS	AI775760	564 bp	mRNA	linear	EST 18-MAY-2000
DEFINITION	EST256860 tomato resistant, Cornell Lycopersicon esculentum cDNA clone CLER16B23, mRNA sequence.				
ACCESSION	AI775760				
VERSION	AI775760.1				
KEYWORDS	GI:5273801				
SOURCE	EST.				
ORGANISM	Lycopersicon esculentum (tomato)				
REFERENCE	Lycopersicon esculentum Euxariota, Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons: core eudicots; asterids; lamids; Solanales; Solanaceae; Solanum; Lycopersicon. D'Ascenzo, M., He, X., Lyman, J., Matern, A.L., Vision, T., Holt, I.E., Liang, F., Upton, J., Ronning, C.M., Craven, M.B., Fujii, C.Y., Bowman, C.L., Nierman, W., Fraser, C.M., Venter, J.C., Tanksley, S.D., Giovannoni, J.J. and Martin, G.B.				
AUTHORS	Generation of ESTs from Pseudomonas resistant tomato				
TITLE	Unpublished				
JOURNAL	Contact: CUGI				
COMMENT	Clemson University Genomics Institute Clemson University 100 Jordan Hall, Clemson, SC 29634, USA Email: <a href="http://www.genome.clemson.edu/orders/index.html">http://www.genome.clemson.edu/orders/index.html</a> 5 prime sequence.				
FEATURES	Location/Qualifiers				
SOURCE	1..564 /organism="Lycopersicon esculentum" /mol_type="mRNA" /cultivar="R10-12 (35S::Pto in Rio Grande x Money Maker)" /db_xref="taxon:4081" /clone="CLER16B23" /tissue_type="leaf" /dev_host="4-week old" /lab_host="SOLR" /clone_lib="tomato resistant, Cornell" /note="Vector: pBluescript SK(-); Site_1: EcoRI; Site_2: XhoI; cLER - Tomato Pseudomonas Resistant EST Library. directionally cloned cDNAs inserted into pBluescript SK( ) at 5' end with EcoRI and 3' end with XhoI site."				
BASE COUNT	134 a 94 c 147 g 189 t				
ORIGIN					
Alignment Scores:					
Pred. No.:	5.2e-55	Length:	564		
Score:	490.00	Matches:	91		
Percent Similarity:	89.66%	Conservative:	13		
Best Local Similarity:	78.45%	Mismatches:	12		
Query Match:	80.72%	Indels:	0		
DB:	9	Gaps:	0		
US-09-786-715-4 (1-118) x AI775760 (1-564)					
OY	3	SerSerGIUGlUGlGlnValIleGlcysHISerValAspClnTrpLysGlnInPhe	22		
Db	21	TCATCTGAGGAAGCAAGTATCGCGCCCAAGAGTGAAGAGTGAAGGCGAGCTC	80		
OY	23	GlnLysGlnValAspSerLysLysLeuValIleAspPheThrAlaSerTrpCysGly	42		
Db	81	CAGAGGGGTGGGACCAAAAACATGGGGTGGGTGGATTACTGCTTCGTGGTGGGT	140		
OY	43	ProCysArgPheIleAlaProIleLeuAlaGluMetAlaLysThrProHisValIle	62		
Db	141	CCCTGGCCGTTTATTATGGCCCAATCTTGTTGGTGCATCTGCAAGAAAGATGCCCATGTTTANG	200		
OY	63	PheLeuLysValAspValAspClnLeuLysThrValAlaGlnGlnPheLysValGluAla	82		
Db	201	TTCTCTCAAGGTGTGATGATCACTGCAAGAAAGTTGCAGAGCAATGCAATGGAGGCA	260		
OY	83	MetProThrPheValPheLeuLysGlnGlnLysGlnValGluArgLeuValGlyAlaArg	102		
Db	261	ATGCCCACTTTGCTTCTTATTAATAAGGGTAAAGAGTGAATGAGGCTTTGTGGCCAAAT	320		

OY 103 LysgluGlueuglnAlaThrValGluLysHisGlyAlaIleThrAla 118  
 Db 321 AAAGACGAGTTGCTTCAGACCATGAGAAAGCATGCTGCTCCTGCT 368  
 RESULT 12  
 AM037831  
 LOCUS AM037831 567 bp mRNA linear EST 18-MAY-2001  
 DEFINITION EST279460 tomato mixed elicitor, BTI Lycopersicon esculentum cDNA  
 clone CLET3L21, mRNA sequence.  
 ACCESSION AM037831  
 VERSION AM037831.1 GI:5896585  
 KEYWORDS EST.  
 SOURCE Lycopersicon esculentum (tomato)  
 ORGANISM Lycopersicon esculentum  
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
 asterids; lamids; Solanales; Solanaceae; Solanum; Lycopersicon.  
 REFERENCE 1 (bases 1 to 567)  
 D'Ascenzo, M., He, X., Lyman, J., Holt, I.E., Liang, F., Upton, J.,  
 Romling, C.M., Craven, M.B., Fujii, C.Y., Bowman, C.L., Nierman, W.,  
 Fraser, C.M., Venter, J.C., Martin, G.B., Tanksley, S.D. and Giovannoni,  
 J.  
 TITLE Generation of ESTs from tomato leaf tissue  
 JOURNAL Unpublished  
 COMMENT Contact: CUGI  
 Clemson University Genomics Institute  
 Clemson University  
 100 Jordan Hall, Clemson, SC 29634, USA  
 Email: <http://www.genome.clemson.edu/orders/index.html>  
 5 prime sequence.  
 FEATURES  
 source Location/Qualifiers  
 1..567  
 /organism="Lycopersicon esculentum"  
 /mol\_type="mRNA"  
 /cultivar="Rio Grande Ptor"  
 /db\_xref="taxon:4081"  
 /clone="CLET3L21"  
 /tissue\_type="leaf"  
 /dev\_stage="4-6 week old plants"  
 /lab\_host="XLI-Blue MRF"  
 /clone\_lib="tomato mixed elicitor, BTI"  
 /note="Vector: pBluescript SK(-); Site.1: EcoRI; Site.2:  
 XhoI; CLET - Inoculated with a variety of disease response  
 elicitors. Plants exposed to 2,6 dichloroisonicotinic  
 acid, BTH, jasmonic acid, ethylene, fenthion, Eix,  
 okadaic acid, or systemin prior to tissue harvest. EcoRI  
 site was destroyed during cloning."  
 BASE COUNT 136 a 87 c 149 g 194 t 1 others  
 ORIGIN  
 Alignment Scores:  
 Pred. No.: 5.24e-55 Length: 567  
 Score: 490.00 Matches: 91  
 Percent Similarity: 89.66% Conservative: 13  
 Best Local Similarity: 78.45% Mismatches: 12  
 Query Match: 80.72% Indels: 0  
 DB: Gaps: 0  
 US-09-786-715-4 (1-118) x AM037831 (1-567)  
 OY 3 SerSerGluGluGlnValIleGlyCysHisSerValAspGluTrpLysGluGlnPhe 22  
 Db 77 TCATCTGAGAGAGACAAAGTGCATCGCTGCCACAGGTTGAGCGAGAGTGCACCTC 136  
 OY 23 GlnLysGlyValAspSerLysLysLeuValValIleAspPheThrAlaSerTrpCysGly 42  
 Db 137 CAGAGGGGTGTGGAGACCAAAAACCTGCTGTGGATTTTACTGCTCTGCTGCGGT 196  
 OY 43 ProCysArpPheIleAlaProIleLeuAlaGluMetAlaLysLysThrProHisValIle 62  
 Db 197 CTTTGGCGTTTATGTCGCCAATTCTTGCGACATTCCTAAGAAGATCCCAATGTATG 256  
 OY 63 PheLeuLysValAspValAspGluLeuLysThrValAlaGluLupPheLysValGluAla 82

Db 257 TTCCTCAAGGTTGATGTTGATGATGAAAGAACTTGACAGAGATGATGTCGAGACA 316  
 OY 83 MetProThrPheValAlaPheLeuLysGluGlyLysGluValGluArgLeuValGlyAlaArg 102  
 Db 317 ATGCCAAGCTTTGCTTCATTAAAGAGGTAAGAAAGGATGAGTTGTTGTTGCCAAT 376  
 OY 103 LysgluGlueuglnAlaThrValGluLysHisGlyAlaIleThrAla 118  
 Db 377 AAAGACGAGTTGCTTCAGACCATGAGAAAGCATGCTGCTCCTGCT 424  
 RESULT 13  
 B1208736  
 LOCUS B1208736 570 bp mRNA linear EST 11-JUL-2001  
 DEFINITION EST526776 cTOS Lycopersicon esculentum cDNA clone cTOS17P22 5' end,  
 mRNA sequence.  
 ACCESSION B1208736  
 VERSION B1208736.1 GI:14686460  
 KEYWORDS EST.  
 SOURCE Lycopersicon esculentum (tomato)  
 ORGANISM Lycopersicon esculentum  
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
 asterids; lamids; Solanales; Solanaceae; Solanum; Lycopersicon.  
 REFERENCE 1 (bases 1 to 570)  
 van der Hoeven, R., Sun, H., Bezzerides, J., Cho, J., Uutterback, R.,  
 Romling, C. and Tanksley, S.  
 TITLE Generation of ESTs from Tomato Suspension Cultures  
 JOURNAL Unpublished  
 COMMENT Contact: CUGI  
 Clemson University Genomics Institute  
 Clemson University  
 100 Jordan Hall, Clemson, SC 29634, USA  
 Email: <http://www.genome.clemson.edu/orders/index.html>  
 FEATURES  
 source Location/Qualifiers  
 1..570  
 /organism="Lycopersicon esculentum"  
 /mol\_type="mRNA"  
 /cultivar="TA496, F6203"  
 /db\_xref="taxon:4081"  
 /clone="cTOS17P22"  
 /tissue\_type="suspension cultures"  
 /lab\_host="SOLR"  
 /clone\_lib="cTOS"  
 /note="Vector: pBluescript SK(-); Site.1: EcoRI; Site.2:  
 XhoI; Suspension cultures of L.esculentum F6203 were grown  
 in Murashige and Skoog based medium, supplemented with 15%  
 coconut milk (filter sterilized and added after  
 autoclaving), 2% saccharose, and 1mg/ml 2,4D (pH5.8).  
 Fresh medium was added every 7 days, and cultures were  
 grown at 25 C, with 12hrs of light and continuous  
 shaking."  
 BASE COUNT 134 a 97 c 148 g 191 t  
 ORIGIN  
 Alignment Scores:  
 Pred. No.: 5.27e-55 Length: 570  
 Score: 490.00 Matches: 91  
 Percent Similarity: 89.66% Conservative: 13  
 Best Local Similarity: 78.45% Mismatches: 12  
 Query Match: 80.72% Indels: 0  
 DB: Gaps: 0  
 US-09-786-715-4 (1-118) x B1208736 (1-570)  
 OY 3 SerSerGluGluGlnValIleGlyCysHisSerValAspGluTrpLysGluGlnPhe 22  
 Db 16 TCATCTGAGAGAGACAAAGTGCATCGCTGCCACAGGTTGAGCGAGAGTGCACCTC 75  
 OY 23 GlnLysGlyValAspSerLysLysLeuValValIleAspPheThrAlaSerTrpCysGly 42  
 Db 76 CAGAGGGGTGTGGAGACCAAAAACCTGCTGTGGATTTTACTGCTCTGCTGCGGT 135

```

OY 43 ProCysArgPheIleAlaProIleLeuAlaIleMetAlaLysLysThrProHisValIle 62
    |||||||
DB 136 CCTGCGGTTTATTTGCCCCCAATTTCTGTGACATTGCTAAGAAAGATGCCCATGTATTAG 195
OY 63 PheLeuLysValAspValAspIleLeuLysThrValAlaIleGluPheLysValGluAla 82
    |||||||
DB 196 TTCCTCAAGGTTGATGTTGATGACACTGAAGAAAGTTGCAGAGCAATGGAATGTGGAGCA 255
OY 83 MetProThrPheValPheLeuLysGluLysGluValGluAlaArgLeuValGluAlaArg 102
    |||||||
DB 256 ATGCCAAGCTTTGTCTCTCACTTAAGAGGGTTAAAGAGTGAATGAGGTGTGTGCTGCCAAT 315
OY 103 LysGluGluLeuGluAlaThrValAlaGluLysHisGluAlaIleThrAla 118
    |||||||
DB 316 AAGACGAGTTCTCTCAGACCATAGAGAAAGCATGTGCTGCTCTCTCT 363

RESULT 14
A1778940 573 bp mRNA linear EST 18-MAY-2001
LOCUS EST259819 tomato susceptible, Cornell Lycopersicon esculentum cDNA
DEFINITION clone CLES6J23, mRNA sequence.
ACCESSION A1778940
VERSION A1778940.1 GI:5276981
KEYWORDS EST.
SOURCE Lycopersicon esculentum (tomato)
ORGANISM Lycopersicon esculentum
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
Asterids; Lamiales; Solanales; Solanales; Solanum; Lycopersicon.
1 (bases 1 to 573)
D'Ascenzo,M., He,X., Lyman,J.J., Matero,A.L., Vision,T., Holt,I.E.,
Liang,F., Upton,J., Ronning,C.M., Craven,M.B., Fujii,C.Y., Bowman
,C.L., Nieman,M., Fraser,C.M., Venter,J.C., Tanksley,S.D.,
Giovannoni,J.J. and Martin,G.B.
Generation of ESTs from Pseudomonas susceptible tomato
Unpublished
JOURNAL Contact: CUGI
COMMENT Clemson University
100 Jordan Hall, Clemson, SC 29634, USA
Email: http://www.genome.clemson.edu/orders/index.html
5 prime sequence.

FEATURES
    source
        1..573
            /organism="Lycopersicon esculentum"
            /mol_type="mRNA"
            /cultivar="R11-13 (Rio Grande x Money Maker)"
            /db_xref="taxon:4081"
            /clone="CLES6J23"
            /tissue_type="leaf"
            /dev_stage="4-week old"
            /lab_host="SOLR"
            /clone_1lb="tomato susceptible, Cornell"
            /note="Vector: pBluescript SK(-); Site,1: EcoRI; Site,2:
            XhoI; CLES - Tomato Pseudomonas Susceptible EST library.
            directionally cloned cDNAs inserted into pBluescript SK(-)
            ) at 5' end with EcoRI and 3' end with XhoI site"

BASE COUNT 137 a 95 c 148 g 193 t

Alignment Scores:
Pred. No.: 5, 31e-55 Length: 573
Score: 490.00 Matches: 91
Percent Similarity: 89.668 Conservative: 13
Best Local Similarity: 78.458 Mismatches: 12
Query Match: 80.728 Indels: 0
Gaps: 0

US-09-786-715-4 (1-118) x A1778940 (1-573)

OY 3 SerSerGluGluGluGluValIleGlyCysHisSerValAspGluTrpLysGluInphe 22
    |||||||
DB 30 TCATCTGAGGAAGACAACTGATCGCTCCACAAAGCTTGAGAGAGTGAAGGTGCAGCTC 89

```

```

OY 23 GluLysGluValAspSerLysLysLeuValIleLeuAspPheThrAlaSerTrpCysGly 42
    |||||||
DB 90 CAGAAAGGTGTGAGAGCCAAAACCTGGGGGTGGGATTTACTGCTTCTGTCGGGT 149
OY 43 ProCysArgPheIleAlaProIleLeuAlaIleMetAlaLysLysThrProHisValIle 62
    |||||||
DB 150 CCTGCGGTTTATTTGCCCCCAATTTCTGTGACATTGCTAAGAAAGATGCCCATGTATTAG 209
OY 63 PheLeuLysValAspValAspIleLeuLysThrValAlaIleGluPheLysValGluAla 82
    |||||||
DB 210 TTCCTCAAGGTTGATGTTGATGACACTGAAGAAAGTTGCAGAGCAATGGAATGTGGAGCA 269
OY 83 MetProThrPheValPheLeuLysGluLysGluValGluAlaArgLeuValGluAlaArg 102
    |||||||
DB 270 ATGCCAAGCTTTGTCTCTCACTTAAGAGGGTTAAAGAGTGAATGAGGTGTGTGCTGCCAAT 329
OY 103 LysGluGluLeuGluAlaThrValAlaGluLysHisGluAlaIleThrAla 118
    |||||||
DB 330 AAGACGAGTTCTCTCAGACCATAGAGAAAGCATGTGCTGCTCTCTCTCT 377

RESULT 15
A1781827 573 bp mRNA linear EST 18-MAY-2001
LOCUS EST262706 tomato susceptible, Cornell Lycopersicon esculentum cDNA
DEFINITION clone CLES17C16, mRNA sequence.
ACCESSION A1781827
VERSION A1781827.1 GI:5279868
KEYWORDS EST.
SOURCE Lycopersicon esculentum (tomato)
ORGANISM Lycopersicon esculentum
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
Asterids; Lamiales; Solanales; Solanales; Solanum; Lycopersicon.
1 (bases 1 to 573)
D'Ascenzo,M., He,X., Lyman,J.J., Matero,A.L., Vision,T., Holt,I.E.,
Liang,F., Upton,J., Ronning,C.M., Craven,M.B., Fujii,C.Y., Bowman
,C.L., Nieman,M., Fraser,C.M., Venter,J.C., Tanksley,S.D.,
Giovannoni,J.J. and Martin,G.B.
Generation of ESTs from Pseudomonas susceptible tomato
Unpublished
JOURNAL Contact: CUGI
COMMENT Clemson University
100 Jordan Hall, Clemson, SC 29634, USA
Email: http://www.genome.clemson.edu/orders/index.html
5 prime sequence.

FEATURES
    source
        1..573
            /organism="Lycopersicon esculentum"
            /mol_type="mRNA"
            /cultivar="R11-13 (Rio Grande x Money Maker)"
            /db_xref="taxon:4081"
            /clone="CLES17C16"
            /tissue_type="leaf"
            /dev_stage="4-week old"
            /lab_host="SOLR"
            /clone_1lb="tomato susceptible, Cornell"
            /note="Vector: pBluescript SK(-); Site,1: EcoRI; Site,2:
            XhoI; CLES - Tomato Pseudomonas Susceptible EST library.
            directionally cloned cDNAs inserted into pBluescript SK(-)
            ) at 5' end with EcoRI and 3' end with XhoI site"

BASE COUNT 137 a 95 c 148 g 193 t

Alignment Scores:
Pred. No.: 5, 31e-55 Length: 573
Score: 490.00 Matches: 91
Percent Similarity: 89.668 Conservative: 13
Best Local Similarity: 78.458 Mismatches: 12
Query Match: 80.728 Indels: 0
Gaps: 0

```

us-09-786-715-4 (1-118) x AT81827 (1-573)

```

QY      3 SerSerGluGluGlnValIleGlyCysHisSerValAspGluTrpLysGluGlnPhe 22
      |||
Db      30 TCATCTGAGGAGGACAGAGTGCCTGCCACAGGTTGAGAGTGAAGTGCACGCTC 89
      |||
QY      23 GlnLysGlyValAspSerLysLysLeuValValIleAspPheThrAlaSerTrpCysGly 42
      |||
Db      90 CAGAGGGGTGAGAGACCAAAACACTGGTGGTGGATTTACTGCTCCTCGGTGGCGT 149
      |||
QY      43 ProCysArgPheIleAlaProIleLeuAlaGluMetAlaLysLysThrProHisValIle 62
      |||
Db      150 CCTTGCCGTTTATTGCCCAATTCCTGCACATGCTAAGAGATGCCCATGTTATG 209
      |||
QY      63 PheLeuLysValAspValAspGluLeuLysThrValAlaGluGluPheLysValGluAla 82
      |||
Db      210 TTCCTCAAGGTTCATGTTGATGAACCTGACAGAAAGTTCCAGAGCAATGATGGAGGCA 269
      |||
QY      83 MetProThrPheValPheLeuLysGluGlyLysGluValGluArgLeuValGlyAlaArg 102
      |||
Db      270 ATGCCAAGCTTTGCTTCATTAAAGAGGTAAGAGAGTGAAGTGGTGGTGGTGGCAAT 329
      |||
QY      103 LysGluGluLeuGlnAlaThrValGluLysHisGlyAlaIleThrAla 118
      |||
Db      330 AAAGACGATTGCTTCAGACCATTAGAGAGCATGTCCTGCTCTCTGCT 377

```

Search completed: August 17, 2003, 23:34:15  
 Job time : 1600.81 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - nucleic search, using frame\_plus\_p2n model

Run on: August 17, 2003, 18:47:50 ; Search time 1856.4 Seconds

(without alignments)  
2644.453 Million cell updates/sec

Title: US-09-786-715-6

Perfect score: 615

Sequence: 1 MAGSSEGGVISCHTVEEMN.....GAKKDELQKIQKHVASASA 120

Scoring table:

BIOSUM62	
Xgapop 10.0 , Xgapext 0.5	
Ygapop 10.0 , Ygapext 0.5	
Fgapop 6.0 , Fgapext 7.0	
Delop 6.0 , Delext 7.0	

Searched: 2888711 seqs, 2045481386 residues

Total number of hits satisfying chosen parameters: 5777422

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+\_p2n.model -DEV=xlp  
-O=/cgn2.1/USPTO/US09786715/runat\_11082003\_150514\_6035/app\_query.fasta\_1.1052  
-DB=GenEmbl -QFMT=fastap -SUFFIX=rge -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0  
-UNITS=bits -START=1 -END=-1 -MATRIX=blsum62 -TRANS=human40.cdi -LIST=45  
-DOCALLIGN=200 -THR.SCORE=pct -THR.MAX=100 -THR.MIN=0 -ALIGN=15 -MODE=LOCAL  
-OUTFMT=ptio -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000  
-USER=US09786715.ecgn\_1.1.5712\_eunat\_11082003\_150514\_6035 -NCPU=6 -ICPU=3  
-NO\_MMAP -LARGEQUERY -NEG.SCORES=0 -WAIT -DSPELCK=100 -LONGLOC  
-DEV.TIMEOUT=120 -WARN.TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6  
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

GenEmbl:\*  
1: gb\_ba:\*  
2: gb\_hcg:\*  
3: gb\_in:\*  
4: gb\_om:\*  
5: gb\_ov:\*  
6: gb\_pat:\*  
7: gb\_ph:\*  
8: gb\_pl:\*  
9: gb\_pt:\*  
10: gb\_ro:\*  
11: gb\_sts:\*  
12: gb\_sy:\*  
13: gb\_un:\*  
14: gb\_vl:\*  
15: em\_ba:\*  
16: em\_fun:\*  
17: em\_hum:\*  
18: em\_in:\*  
19: em\_mu:\*  
20: em\_om:\*  
21: em\_or:\*  
22: em\_ov:\*  
23: em\_pat:\*  
24: em\_ph:\*  
25: em\_pl:\*  
26: em\_ro:\*  
27: em\_sts:\*  
28: em\_un:\*

29: em\_vl:\*  
30: em\_hcg\_hum:\*  
31: em\_hcg\_in:\*  
32: em\_hcg\_other:\*  
33: em\_hcg\_mus:\*  
34: em\_hcg\_pln:\*  
35: em\_hcg\_rod:\*  
36: em\_hcg\_mam:\*  
37: em\_hcg\_vrt:\*  
38: em\_sy:\*  
39: em\_hlgo\_hum:\*  
40: em\_hlgo\_mus:\*  
41: em\_hlgo\_other:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	544	88.5	630	PSA310990	AJ310990 Pismum sat
2	490	79.7	603	RCPHIORXN	Z70677 R.communis
3	482	78.4	345	AX505468	AX505468 Sequence
4	482	78.4	480	AY088687	AY088687 Arabidops
5	482	78.4	497	ATTHTIOARA	Z14084 A.thaliana
6	460	74.8	664	AF323593	AF323593 Prunus pe
7	456	74.1	698	NTTFRNA	X58527 N.tabacum m
8	453	73.7	665	AY170650	AY170650 Pismum sat
9	445	72.4	653	AR016869	AR016869 Sequence
10	445	72.4	653	AR020895	AR020895 Sequence
11	445	72.4	653	AR027218	AR027218 Sequence
12	445	72.4	653	AR038505	AR038505 Sequence
13	445	72.4	653	AR064647	AR064647 Sequence
14	445	72.4	653	AR067572	AR067572 Sequence
15	445	72.4	653	I38524	I38524 Sequence 10
16	445	72.4	653	I56999	I56999 Sequence 10
17	445	72.4	653	I59865	I59865 Sequence 10
18	445	72.4	653	I75192	I75192 Sequence 10
19	434	70.6	784	AY271308	AY271308 Citrus x
20	429	69.8	357	AY040028	AY040028 Arabidops
21	429	69.8	528	ATTHRED4	Z35476 A.thaliana
22	429	69.8	530	AK118035	AK118035 Arabidops
23	429	69.8	560	AY087159	AY087159 Arabidops
24	429	69.8	561	AF360227	AF360227 Arabidops
25	414	67.3	593	PSA319808	AJ319808 Pismum sat
26	399	64.9	345	AF483265	AF483265 Populus t
27	388	63.1	526	G73679	G73679 RZ488R etio
28	388	63.1	601	OSG92541	U92541 Oryza sativ
29	388	63.1	686	E08194	E08194 RPS13 gene
30	388	63.1	687	RICHTH	D21836 Oryza sativ
31	385.5	62.7	590	AY088698	AY088698 Arabidops
32	385	62.6	595	D87984	D87984 Fagopyrum e
33	383	62.3	366	AX654096	AX654096 Sequence
34	381.5	62.0	360	AX505469	AX505469 Sequence
35	381.5	62.0	360	BT004710	BT004710 Arabidops
36	381.5	62.0	556	AK118542	AK118542 Arabidops
37	381	62.0	740	AB053294	AB053294 Oryza sat
38	378.5	61.5	642	ATTHRED1	Z35473 A.thaliana
39	376.5	61.2	357	AX412276	AX412276 Sequence
40	376.5	61.2	357	AX412277	AX412277 Sequence
41	376.5	61.2	357	AX412278	AX412278 Sequence
42	376.5	61.2	357	AX412279	AX412279 Sequence
43	376.5	61.2	357	AX412473	AX412473 Sequence
44	376.5	61.2	357	AX412474	AX412474 Sequence
45	376.5	61.2	357	AX412566	AX412566 Sequence

RESULT 1

#### ALIGNMENTS

LOCUS	PSA310990	630 bp	mRNA	linear	PLN 12-APR-2001
DEFINITION	PSA310990 Pisum sativum mRNA for thioredoxin h (trxh gene).				
ACCESSION	AJ310990				
VERSION	AJ310990.1 GI:1362483				
KEYWORDS	thioredoxin H; trxh gene.				
SOURCE	Pisum sativum (pea)				
ORGANISM	Pisum sativum				
	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids; eurosids I; Fabales; Fabaceae; Papilionoideae; Viciaeae; Pisum.				
REFERENCE	1 Traverso, J.A. Thesis (2001) Department of Bioquimica, Biologia Celular y Molecular de Plantas, Estacion Experimental del Zaidin (C.S.I.C.), Granada, Spain				
AUTHORS	2 Traverso, J.A., Gazalis, R., Sahrawy, M., Lopez-Gorge, J. and Chueca, A.				
TITLE	Sequencing, cloning and expression of Pisum sativum Thioredoxin h				
JOURNAL	Unpublished				
REFERENCE	3 (bases 1 to 630)				
AUTHORS	Traverso, J.A.				
TITLE	Direct Submission				
JOURNAL	Submitted (11-APR-2001) Traverso J.A., Biologia Celular y Molecular de plantas, Estacion Experimental del Zaidin (CSIC), Profesor Albareda I, Granada, 18008, SPAIN				
FEATURES	Location/Qualifiers				
source	1..630				
	/organism="Pisum sativum"				
	/mol_type="mRNA"				
	/variety="Lincoln"				
	/db_xref="taxon:3888"				
	/tissue_type="Leaves"				
gene	1..630				
	/gene="trxh"				
CDS	36..398				
	/gene="trxh"				
	/function="Unknown"				
	/codon_start=1				
	/evidence="experimental"				
	/product="thioredoxin h"				
	/protein_id="CAC36986.1"				
	/db_xref="GI:1362488"				
	/translation="MAGSSSEGGVVISCHSDAMDNDLHNGNSKLLIVDFTASGCP				
	CRFLAPLIGELAKKFTNVIPLKVDYDELKSYADQMAVEAMPTFVKEEGTILKRYVCA				
	KKEELOOTIERHVASNA"				
3'UTR	398..609				
	/gene="trxh"				
polyA_site	609				
	/gene="trxn"				
BASE COUNT	195 a	95 c	141 g	199 t	
ORIGIN					
Alignment Scores:					
Pred. No.:	4.38e-55	Length:		630	
Score:	544.00	Matches:		102	
Percent Similarity:	95.00%	Conservative:		12	
Best Local Similarity:	85.00%	Mismatch:		6	
Query Match:	88.46%	Indels:		0	
DB:	8	Gaps:		0	
US-09-786-715-6 (1-120) x PSA310990 (1-630)					
OY	1	MetlaaglyserSerGluGluGlyGlnValIleSerCysHisThrValGluGluTyrAsn 20			
Db	36	ATGCGACGTTTCATAGAGAGGACAGTTCACGCTGTACACGCGTTGATGATGAGAAC 95			
OY	21	AspGlnLeuGlnIlysgIyAsnGluSerLysLysLeuIleValValaAspPheThrAlaSer 40			
Db	96	GATATCCTTCACAGAGCAATGATTCACAGAACTGATTTGTGTGGACATTACTGCTTCT 155			
OY	41	TTPCysGlyProCysArgPheIleAlaIleProPheLeuAlaGluLeuAlaLysLysPheThr 60			

```

Db      156 TGGTGTGACCATCGCCGTTTCAATTCGACCATTCCTTGSTGATTTGGCTTAAGAGTTTACA 215
QY      61 SerValIlePheLeuLysValAspValAspGluLeuLysSerValSerGlnAspTrpPala 80
      ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
Db      216 AATGCATATTCTCTTAAGGTGAGGTGACGACGAACTAAAGTCTGTTGCTCAAGATTGGGCT 275
QY      81 IleGluAlaMetProThrPheValPheValLysGluGlyTrpLeuLeuAspLysValVal 100
      ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
Db      276 GTTGGGCTATGCCAACATTGTTGCTTTGTAAGAAGAACAGCATTTGGCAAACTGGT 335
QY      101 GlyAlaLysAspGluLeuGlnGlnLysIleGlnLysSHSLysValAlaSerAlaSerAla 120
      |||||::::::::::::::::::::::::::::| |:::::|||||:::::|||||:::::||||
Db      336 GGAGCAAGAAAGAAAGAACTGCAGCAGACATTCGAACATGTGCTTCACTATATGCT 395

RESULT 2
RCHTIORXN      603 bp      mRNA      linear      PLN 04-APR-1996
LOCUS      R.communis mRNA for thioredoxin.
ACCESSION      Z70677
VERSION      Z70677.1 GI:1255953
KEYWORDS      thioredoxin.
SOURCE      Ricinus communis (castor bean)
ORGANISM      Ricinus communis
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
rosids; eurosids I; Malpighiales; Euphorbiaceae; Ricinus.
REFERENCE      1 (bases 1 to 603)
AUTHORS      Szederkenyi,J., Dolgenier,E. and Schoberl,C.
TITLE      Szederkenyi,J., Dolgenier,E. and Schoberl,C.
JOURNAL      cDNA expressed in Ricinus cotyledons
REFERENCE      2 (bases 1 to 603)
AUTHORS      Dolgenier,E.
TITLE      Direct Submission
JOURNAL      Submitted (04-APR-1996) Dolgenier E., University of Bayreuth,
      Department of Plant Physiology, Universitaetsstr.30, Bayreuth,
      Germany, 95440

FEATURES
    source
        1..603
            /organism="Ricinus communis"
            /mol_type="mRNA"
            /cultivar="Sanguineus"
            /db_xref="taxon:3988"
            /clone="pEDRH018"
            /tissue_type="cotyledon"
            /clone_lib="lambda Excell"
            /dev_stage="seedling"
            23..379
            /codon_start=1
            /product="thioredoxin"
            /protein_id="CAA94534.1"
            /db_xref="GI:1255954"
            /db_xref="SPTREMBL:O43636"
            /translation="MAAEQGVICSHYEAANEOLOGNDTKGLIVDFATSMWGPCPCG
            FIAPFLAELAKLEPNVTPFLKVDVDELKTVAHENAVESPTFMFLKEGIMDKYVGAKKK
            DELQQTIAKHMAVST"

BASE COUNT      179 a      101 c      157 g      166 t

ORIGIN
Alignment Scores:
Pred. No.:      1.01e-48      Length:      603
Score:      490.00      Matches:      89
Percent Similarity:      91.38%      Conservative:      17
Best local Similarity:      76.72%      Mismatches:      10
Query Match:      79.67%      Indels:      0
DB:      8      Gaps:      0

US-09-786-715-6 (1-120) x RCHTIORXN (1-603)
QY      4 SerSerGluGluGluGlnValIleLeSerCysHisrThrValGluGluTrpAsnAspGlnLeu 23
      ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
Db      26 GCACACAGAAAGGCGCAAGTGATCGGTGCGCACACTGTTGAGCATGATGAGCAATTTG 85

```

OY 24 GlnysGlyAsnGluSerIysLysLeuIleValIAspPheThrAlaSerTrpCysGly 43  
 |||||.....||| |||||.....|||  
 Db 86 CAGAGGGAATGATACCAAGGACGATCGTTGATTTACTCTTCATCGTGTGA 145  
 OY 44 ProCysArgPheIleAlaProPheLeuAlaGluLeuAlaLysLysPheThrSerValIle 63  
 |||||.....||| |||||.....|||  
 Db 146 CCATGCGCGTTTCATGCTCCCTTCTTGCTGAGCTGGCCAGAACTGCCAATGTTACC 205.  
 OY 64 PheLeuLysValAspValAspLeuLysSerValSerGlnAspTrpAlaIleGluAla 83  
 |||||.....||| |||||.....|||  
 Db 206 TTCCTGGAAGTGAATGTGATGACGAGACTGTCACAGAGCGCTGTGGAGTCA 265  
 OY 84 MetProthPheValPheValLysGluGlyThrLeuLeuAspLysValValGlyAlaLys 103  
 |||||.....||| |||||.....|||  
 Db 266 ATGCCCACTTATATGTTCTCTGAAGGAGGGAAGATTATGACACAGCTGTGGAGCAAG 325  
 OY 104 LysAspGluLeuGlnGlnLysIleGlnLysHisValAlaSerAlaSer 119  
 |||||.....||| |||||.....|||  
 Db 326 AAGAGCAAGACTGCAGCAACTATAGCGAACAACATGCTACTGCTCC 373  
 RESULT 3  
 AX505468  
 LOCUS AX505468 345 bp DNA linear PAT 27-SEP-2002  
 DEFINITION Sequence 163 from Patent WO0216655.  
 ACCESSION AX505468  
 VERSION AX505468.1 GI:23386705  
 KEYWORDS  
 SOURCE Arabidopsis thaliana (thale cress)  
 ORGANISM Arabidopsis thaliana  
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
 rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsids.  
 REFERENCE 1  
 AUTHORS Harper, J.F., Kreps, J., Wang, X. and Zhu, T.  
 TITLE Stress-regulated genes of plants, transgenic plants containing  
 same, and methods of use  
 JOURNAL Patent: WO 0216655-A 163 28-FEB-2002;  
 The Scripps Research Institute (US); Syngenta Participations AG  
 (CH)  
 FEATURES  
 source location/Qualifiers  
 1..345  
 /organism="Arabidopsis thaliana"  
 /mol\_type="genomic DNA"  
 /db\_xref="taxon:3702"  
 BASE COUNT 90 a 68 c 90 g 97 t  
 ORIGIN  
 Alignment Scores:  
 Pred. No.: 4.65e-48 Length: 345  
 Score: 482.00 Matches: 88  
 Percent Similarity: 89.38% Conservative: 13  
 Best Local Similarity: 77.88% Mismatches: 12  
 Query Match: 78.37% Indels: 0  
 Gaps: 0  
 US-09-786-715-6 (1-120) x AX505468 (1-345)  
 OY 4 SerSerGluGluGlyGlnValIleSerCysHisThrValGluGluTrpAsnAspGlnLeu 23  
 |||||.....||| |||||.....|||  
 Db 4 GCTTCGGAAGAGGACAACTGATCGCTCCACACCGCTGAGACATGGAACGACGCTT 63  
 OY 24 GlnysGlyAsnGluSerIysLysLeuIleValIAspPheThrAlaSerTrpCysGly 43  
 |||||.....||| |||||.....|||  
 Db 64 CAGAAAGCTAATGATCCAAACTCTTGTGGTGGTGGTTCACGGCTTCTTGGTGGTGA 123  
 OY 44 ProCysArgPheIleAlaProPheLeuAlaGluLeuAlaLysLysPheThrSerValIle 63  
 |||||.....||| |||||.....|||  
 Db 124 CCATGCGCTTCATGCTCCCTCATTTCTGCTGATTTGCGTAAGAACTCCTCAACGTCCT 183  
 OY 64 PheLeuLysValAspValAspLeuLysSerValSerGlnAspTrpAlaIleGluAla 83  
 |||||.....||| |||||.....|||  
 Db 184 TTCCTGGAAGTGAATGTGATGAGTGGTGGCAAGATGATGGCGATACAGCGC 243

OY 84 MetProthPheValPheValLysGluGlyThrLeuLeuAspLysValValGlyAlaLys 103  
 |||||.....||| |||||.....|||  
 Db 244 ATGCCCACTTATATGTTCTCTGAAGGAGGGAAGATTATGACACAACTGTTGGAGCCAA 303  
 OY 104 LysAspGluLeuGlnGlnLysIleGlnLysHisValAla 116  
 |||||.....||| |||||.....|||  
 Db 304 AAGATGAGCTTCAGCTCACTGACATTCGCCAACAACACTGGCT 342  
 RESULT 4  
 AY088687  
 LOCUS AY088687 480 bp mRNA linear PLN 14-APR-2003  
 DEFINITION Arabidopsis thaliana clone 9189 mRNA, complete sequence.  
 ACCESSION AY088687  
 VERSION AY088687.1 GI:21407461  
 KEYWORDS  
 SOURCE Arabidopsis thaliana (thale cress)  
 ORGANISM Arabidopsis thaliana  
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
 rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsids.  
 REFERENCE 1 (bases 1 to 480)  
 AUTHORS Haas, B.-J., Volfovsky, N., Town, C.D., Troukhan, M., Alexandrov, N.,  
 Feldmann, K.A., Flavell, R.B., White, O. and Salzberg, S.L.  
 TITLE Full-length messenger RNA sequences greatly improve genome  
 annotation  
 JOURNAL Genome Biol. 3 (6), RESEARCH0029 (2002)  
 MEDLINE 22088475  
 PUBMED 12093376  
 REFERENCE 2 (bases 1 to 480)  
 AUTHORS Brover, V., Troukhan, M., Alexandrov, N., Lu, Y.-P., Flavell, R. and  
 Feldmann, K.  
 TITLE Full-length cDNA from Arabidopsis thaliana  
 JOURNAL Unpublished  
 3 (bases 1 to 480)  
 REFERENCE 3 (bases 1 to 480)  
 AUTHORS Brover, V., Troukhan, M., Alexandrov, N., Lu, Y.-P., Flavell, R. and  
 Feldmann, K.  
 TITLE Direct Submission  
 JOURNAL Submitted (11-MAR-2002) Ceres, Inc. 3007 Malibu Canyon Road,  
 Malibu, CA 90265, USA  
 COMMENT  
 This clone sequence is one of 5,000 Ceres full-length cDNAs made  
 available to TIGR and Genbank. The following quality assessment of  
 this set was done by comparison with known proteins: two percent of  
 the clones are estimated to be 5'-truncated; less than one percent  
 are 3'-truncated; approximately two percent represent alternative  
 splice variants, including unspliced introns and spliced exons; one  
 percent may contain premature stop codons; five percent may have  
 frame shifts in a coding region. A sequence is considered to be  
 5'-truncated if it lacks the translation initiation start (ATG). A  
 sequence is considered to be 3'-truncated if it lacks the  
 C-terminal end of the encoded protein. Please note that these cDNA  
 sequences are derived from the WS or Laer ecotypes and therefore  
 may contain polymorphisms when compared to sequences from Col-0.  
 Geneset carried out the library production and sequencing of the  
 full-length clones. Ceres, Inc. carried out the clustering of the  
 5' sequences, selection of clones, and sequence assembly.  
 FEATURES  
 source location/Qualifiers  
 1..480  
 /organism="Arabidopsis thaliana"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:3702"  
 /clone="9189"  
 13..357  
 /codon\_start=1  
 /product="thioredoxin h"  
 /protein\_id="AA067008.1"  
 /db\_xref="GI:21617958"  
 /translation="MASSEGOYIACHTVETMNEQLOKANEKTLVVDTASMCPCR  
 FLAPFADIAKKLPVLFKVDTEKLSVASQWAIQADPTFFFLKRGKILDRVVGAKK  
 DELQSTIAKHA"  
 BASE COUNT 124 a 93 c 111 g 152 t  
 ORIGIN  
 Alignment Scores:





/protein\_id="AL26915.1"  
 /db\_xref="GI:1658843"  
 /translation="MAEPNOVIGCHTTOAMEQOLHKNGENKKLVVDFPASMCGPCRL  
 IAPIILAEKKTPEYFELKVDVDELTVSEKGVEMPTFLFLKEKIDKRVGAKKD  
 ELQIKYAKHVAHAASATASATASATATATATATA"

BASE COUNT 182 a 147 c 161 g 174 t

ORIGIN

## Alignment Scores:

Pred. No.: 3,98e-45 Length: 664  
 Score: 460.00 Matches: 83  
 Percent Similarity: 86.21% Conservative: 17  
 Best Local Similarity: 71.55% Mismatches: 16  
 Query Match: 74.80% Indels: 0  
 DB: Gaps: 0

US-09-786-715-6 (1-120) x AF323593 (1-664)

QY 5 SerGluGluGlyGlnValIleSerCysHisThrValGluGluTrpAsnAspGlnLeuGln 24  
 :||||| :||||| :||||| :||| :|||||  
 Db 74 GCGGAGAAATCAAGTCAATCGGCTGCCACACACAGCCCGGAGAGAGACGCTCCAT 133  
 QY 25 LysGlyAsnGluSerLysLysLeuIleValAlaAspPheThrAlaSerTrpCysGlyPro 44  
 :||||| :||||| :||||| :||||| :|||||  
 Db 134 AAGGAAACGAGAAACAAGAACTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 193  
 QY 45 CysArgPheIleAlaProPheLeuAlaGluLeuAlaLysLysPheThrSerValIlePhe 64  
 :||||| :||||| :||||| :||||| :|||||  
 Db 194 TGCCGGTTCATCCGCCCAATCTTGGCGAGTTCGGCTAGAGAGAGAGAGAGAGAGAG 223  
 QY 65 LeuLysValAspValAspGluLeuLysSerValSerGlnAspTrpAlaIleGluAlaMet 84  
 :||||| :||||| :||||| :||||| :|||||  
 Db 254 CTAAAGCTGGACGTGATGATGACGTGATGATGATGATGATGATGATGATGATGATG 313  
 QY 85 ProThrPheValPheValLysGluGlyThrLeuLeuAspLysValValGlyAlaLysLys 104  
 :||||| :||||| :||||| :||||| :|||||  
 Db 314 CCTACCTCTCTCTCTCTCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 373  
 QY 105 AspGluLeuGlnGlnIleGlnLysHisValAlaSerAlaSerAla 120  
 :||||| :||||| :||||| :||||| :|||||  
 Db 374 GACGATTCACATCAAGTTCAGCAAGCATGTGCGCGCGCGCT 421

## RESULT 7

## NTTRNA

LOCUS NTTRNA 698 bp mRNA linear PLN 15-FEB-1994  
 DEFINITION N. tabacum mRNA for thioredoxin.  
 ACCESSION X58527.1 GI:20046  
 VERSION X58527.1  
 KEYWORDS thioredoxin.

SOURCE Nicotiana tabacum (common tobacco)  
 ORGANISM Nicotiana tabacum

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
 Asteroidei; Lamiales; Solanales; Solanaceae; Nicotiana.

REFERENCE 1 (bases 1 to 698)  
 Marty, I. and Meyer, Y.

TITLE Nucleotide sequence of a cDNA encoding a tobacco thioredoxin  
 JOURNAL Plant Mol. Biol. 17 (1), 143-147 (1991)

## MEDLINE

## PUBMED

1868216 2 (bases 1 to 698)

## REFERENCE

## AUTHORS

## TITLES

## JOURNAL

## MEDLINE

## PUBMED

## REFERENCE

## AUTHORS

## TITLES

## JOURNAL

## MEDLINE

## PUBMED

## REFERENCE

## AUTHORS

## TITLES

COMMENT Gene product is probably cytoplasmic.  
 FEATURES Location/Qualifiers  
 source 1..698  
 /organism="Nicotiana tabacum"  
 /mol\_type="mRNA"  
 /strain="white Burley"  
 /db\_xref="taxon:4097"  
 /clone\_lib="plasmid pTZ19"  
 /dev\_stage="in vitro cells"  
 /evidence="experimental"

## mRNA

1..682  
 /evidence="experimental"

## CDS

## CDS

## CDS

## CDS

## CDS

## CDS

## CDS

## CDS

## CDS

## CDS

## CDS

## CDS

## CDS

## CDS

## CDS

## CDS

## CDS

## CDS

## CDS

## CDS

## CDS

## CDS

## CDS

## CDS

## CDS

## CDS

## CDS

## CDS

## CDS

## CDS

## CDS

## CDS

## CDS

## CDS

## CDS

## CDS

## CDS

## CDS

## CDS

## CDS

## CDS

## CDS

## CDS

## CDS

## CDS

## CDS

## CDS

## CDS

## CDS

## CDS

## CDS

## CDS

## CDS

## CDS

## CDS

## CDS

## CDS

US-09-786-715-6 (1-120) x NTTRNA (1-698)

QY 2 AlaGlySerSerGluGluGlnValIleSerCysHisThrValGluGluTrpAsnAsp 21  
 :||| :||||| :||||| :||||| :|||||  
 Db 94 GCTACTTCATCCAGCAGGAGGACAGTCTTCCGCTGCCACAGGTTGAGGATGAGAGAG 153  
 QY 22 GlnLeuGlnLysGlnLysSerLysLysLeuIleValAlaAspPheThrAlaSerTrp 41  
 :||||| :||||| :||||| :||||| :|||||  
 Db 154 TACTTCAGAGAGAGCGGTGAGACTAGAGAACTGGTGGTGGTGGTGGTGGTGGTGG 213  
 QY 42 CysGlyProCysArgPheIleAlaProPheLeuAlaGluLeuAlaLysLysPheThrSer 61  
 :||||| :||||| :||||| :||||| :|||||  
 Db 214 TGCGGCCCTTCCGCTTATTTGCCCCCAATTCCTTGATGACATTTGCTAAGAGATGCC 273  
 QY 62 ValIlePheLeuLysValAspValAspGluLeuLysSerValSerGlnAspTrpAlaIle 81  
 :||||| :||||| :||||| :||||| :|||||  
 Db 274 GTTATATTCCTCAAGTGTGATGTGATGAACCTGAACAGCTTTCCACCGGATGAGCTGC 333  
 QY 82 GluAlaMetProThrPheValPheValLysGluGlyThrLeuLeuAspLysValValGly 101  
 :||||| :||||| :||||| :||||| :|||||  
 Db 334 GAGCAATGCGCACTTTTCTTCATTAAAGATGAGAAAGATGAGACAGATGTTGTGT 393  
 QY 102 AlaLysLysAspGluLeuGlnGlnLysLysLysLysHisValAlaSerAlaSer 119  
 :||||| :||||| :||||| :||||| :|||||  
 Db 394 GCCAAGAAAGAGAGATTGACAGACCATAGTGAAGCATGCTCTCTCTACT 447

## Alignment Scores:

Pred. No.: 1.25e-44 Length: 698  
 Score: 456.00 Matches: 84  
 Percent Similarity: 84.75% Conservative: 16  
 Best Local Similarity: 71.19% Mismatches: 18  
 Query Match: 74.15% Indels: 0  
 DB: Gaps: 0

polysignal  
 polysignal  
 polysignal  
 BASE COUNT 192 a 117 c 164 g 225 t

## RESULT 8

## LOCUS

## DEFINITION

## ACCESSION

## VERSION

## KEYWORDS

## SOURCE

## ORGANISM

## TITLES

## JOURNAL

## MEDLINE

## PUBMED

## REFERENCE

## AUTHORS

## TITLES

## JOURNAL

AY170650 665 bp mRNA linear PLN 02-JAN-2003  
 DEFINITION Pisum sativum clone 3 thioredoxin h mRNA, complete cds.  
 ACCESSION AY170650  
 VERSION AY170650.1 GI:27466893  
 KEYWORDS Pisum sativum (pea)  
 SOURCE Pisum sativum  
 ORGANISM Pisum sativum

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
 Rosids; eustosids I; Fabales; Fabaceae; Papilionoideae; Viciaeae;  
 Pisum.



Oy	2	AlaGlySerSerGIuGIuGIuGIuValIleSerCySHSThrValGIuGIuTrpAsn	21
Db	63	GCTACTTCATCCGAGGAGGAGCAGGTGTCCGCTGCCCAAGATTGAGAAATGACGAG	122
Oy	22	GlnGluGlnGlyAsnGluSerLysLysLeuIleValValAspPheThrAlaSerTrp	41
	:::		
Db	123	TACTTCAAGAAAGCGTTGAGACTAAGAACTGGTGGTGCATTTTACGTCTCATGG	182
Oy	42	CysGlyProCysAlaGpHeIleAlaProPheLeuAlaGluLeuAlaLysPheThrSer	61
Db	183	TGCGGSCCTGGCCGTTTATTTATGGCCCAATTCCTGCAATTCCTAAGAAATGCCCAT	242
Oy	62	ValIlePheLeuLysValAspValAspGluLeuLysSerValSer-GlnAspTrpAlaIle	81
Db	243	GTTATATTCCCAAGGTTGATGTTGATGATGACTGAAGACTGTTTACGCGGAATGAGCTGT	302
Oy	81	eGluAlaMetProThrPheValPheValLysGluGlyThrLeuLeuAspLysValAlaG1	101
Db	303	GGAGGCAATGCGCACTTTTGCTCTCATTAAGATGAGAAAGAAATGGACAGATGTGTGG	362
Oy	101	ValAlaLysLysAspLueGluGlnGlnLysIleGlnLysHisValAlaSerAlaSer	119
Db	363	TGCCAAGAAAGAGGATGGCAGCAGACCATATGAGACATGCTCTCTGCTACT	417
RESULT 11			
AR027218		653 bp	DNA
LOCUS	AR027218		linear
DEFINITION	Sequence 103 from patent US 5856154.		PAT 29-SEP-1999
ACCESSION	AR027218		
VERSION	AR027218.1		GI:5938058
KEYWORDS			
SOURCE	Unknown.		
ORGANISM	Unclassified.		
REFERENCE	1 (bases 1 to 653)		
AUTHORS	Ryals,J.A., Alexander,D.C., Goodman,R.M. and Ward,E.R.		
TITLE	Method of protecting plants from comycete pathogens		
JOURNAL	Patent: US 5856154-A 103 05-JAN-1999;		
FEATURES	location/Qualifiers		
source	1..653		
	/organism="unknown"		
BASE COUNT	165 a 114 c 157 g 215 t 2 others		
ORIGIN			
Alignment Scores:			
Pred. No.:	2.32e-43	Length:	653
Score:	445.00	Matches:	84
Percent Similarity:	84.87%	Conservative:	17
Best Local Similarity:	70.59%	Mismatches:	17
Query Match:	72.36%	Indels:	1
DB:	6	Gaps:	0
US-09-786-715-6 (1-120) x AR027218 (1-653)			
Oy	2	AlaGlySerSerGIuGIuGIuGIuValIleSerCySHSThrValGIuGIuTrpAsn	21
Db	63	GCTACTTCATCCGAGGAGGAGCAGGTGTCCGCTGCCCAAGATTGAGAAATGACGAG	122
Oy	22	GlnGluGlnGlyAsnGluSerLysLysLeuIleValValAspPheThrAlaSerTrp	41
	:::		
Db	123	TACTTCAAGAAAGCGTTGAGACTAAGAACTGGTGGTGCATTTTACGTCTCATGG	182
Oy	42	CysGlyProCysAlaGpHeIleAlaProPheLeuAlaGluLeuAlaLysPheThrSer	61
Db	183	TGCGGSCCTGGCCGTTTATTTATGGCCCAATTCCTGCAATTCCTAAGAAATGCCCAT	242
Oy	62	ValIlePheLeuLysValAspValAspGluLeuLysSerValSer-GlnAspTrpAlaIle	81
Db	243	GTTATATTCCCAAGGTTGATGTTGATGATGACTGAAGACTGTTTACGCGGAATGAGCTGT	302
Oy	81	eGluAlaMetProThrPheValPheValLysGluGlyThrLeuLeuAspLysValAlaG1	101

Db	303	GGAGGCAATGCCA	CTTTGCTTCATTAAGATGAAAAAGAGTGGACAGAGTTGTTGG	362	
Qy	101	YALALysLysAsp	LeuGlnGlnLysIleGlnLysHisValAlaSerAlaSer	119	
Db	363	TGCCAAGAAAGAGAGT	GTTCGACGACCATATGATGAAGATGCTCCTGCTACT	417	
RESULT 12					
AR038505	LOCUS	AR038505	653 bp	DNA	
DEFINITION	Sequence 103 from patent US 5804693.		linear	PAT 29-SEP-1999	
ACCESSION	AR038505				
VERSION	AR038505.1	GI:5957222			
KEYWORDS					
SOURCE	Unknown.				
ORGANISM	Unknown.				
REFERENCE	Unclassified.				
AUTHORS	1 (bases 1 to 653) Gaffney,T.D., Ryals,J.A., Friedrich,L.B., Uknes,S.J., Ward,E.R., Keesmann,H. and Vernoolt,J.B.T.				
TITLE	Chemically regulatable and anti-pathogenic DNA sequences and uses thereof				
JOURNAL	Patent: US 5804693-A 103 08-SEP-1998;				
FEATURES	Location/Qualifiers				
source	1..653				
BASE COUNT	165 a 114 c 157 g 215 t		2 others		
ORIGIN	/organism="unknown"				
Alignment Scores:					
Pred. No.:	2.32e-43	Length:	653		
Score:	445.00	Matches:	84		
Percent Similarity:	84.87%	Conservative:	17		
Best Local Similarity:	70.59%	Mismatches:	17		
Query Match:	72.36%	Indels:	1		
DB:	6	Gaps:	0		
US-09-786-715-6 (1-120) x AR038505 (1-653)					
Qy	2	AlaGlySerSerGlnGlnGlnValIleSerCysHisThrValGlnGluTrpAsnAsp	21		
Db	63	GCTACTTCATCCGAGGAGGAGGACAGTGTGGCGTGGCCACAAGTTGAGGATGGACGAG	122		
Qy	22	GlnGlnGlnGlyAsnGlnSerLysLysLeuIleValValAspPheThrAlaSerTrp	41		
Db	123	TACTTCAGAAAGCGTTGGACACTAAGAACTGGTGGCGATTTTACTGCTTCATGG	182		
Qy	42	CysGlnProCysArgPheIleAlaProPheLeuAlaGlnLeuAlaLysPheThrSer	61		
Db	183	TGCGGSCCTTCCGCTTTATTATGGCCCAATTTCTCTGACATTTGCTAAAGATGCCCAT	242		
Qy	62	ValIlePheLeuLysValAspValAspGlnLeuLysSerValSer-GlnAspTrpAlaIle	81		
Db	243	GTTATATTTCCTCAAGCTTGATGTTGATGAACCTGTTTACGGCGAATGGAGTCT	302		
Qy	81	eGlnAlaMetProThrPheValPheValLysGlnGlyThrLeuLeuAspLysValIleG	101		
Db	303	GGAAGCATGCACTTTTGCTTCATTAAGATGAAAAAGAGTGGACAGAGTGTGG	362		
Qy	101	YALALysLysAspLeuGlnGlnLysIleGlnLysHisValAlaSerAlaSer	119		
Db	363	TGCCAAGAAAGAGAGT	GTTCGACGACCATATGATGAAGATGCTCCTGCTACT	417	
RESULT 13					
AR064647	LOCUS	AR064647	653 bp	DNA	
DEFINITION	Sequence 103 from patent US 5847258.		linear	PAT 29-SEP-1999	
ACCESSION	AR064647				
VERSION	AR064647.1	GI:5993955			
KEYWORDS					
SOURCE	Unknown.				
ORGANISM	Unknown.				
REFERENCE	Unclassified.				
	1 (bases 1 to 653)				



```

Qy      81  egluaIametProthrPheValPheValIysGluGlyThrLeuAspLysValValGI 101
      : ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db      303  GAGGCAATGCCAAGCTTTGCTCTTCATTAAGATGAGAAAGAGAGTGGACAGAGTTGTTGG 362

Qy      101  yAlaIysIysAspGluLeuGlnGlnIysIleGlnIysHisValAlaSerAlaSer 119
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db      363  TGCAGAGAAAGAGAGTTCAGCAGACATAGTGAAGCATGCTGCTGCTACT 417

```

Search completed: August 17, 2003, 21:41:56  
 Job time : 1858.4 secs



GenCore version 5.1.6  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - nucleic search, using frame\_plus\_p2n model

Run on: August 17, 2003, 17:46:53 ; Search time 168.696 seconds  
(without alignments)  
1920.219 Million cell updates/sec

Title: US-09-786-715-6  
Perfect score: 615  
Sequence: 1 MAGSSEGGVISCHTVEEMN.....GAKKDELOQIKKHVASASA 120

Scoring table:  
BLOSUM62  
Xgapop 10.0, Xgapext 0.5  
Ygapop 10.0, Ygapext 0.5  
Fgapop 6.0, Fgapext 7.0  
Delop 6.0, Delext 7.0

Searched: 2552756 seqs, 1349719017 residues

Total number of hits satisfying chosen parameters: 5105512

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Command line parameters:  
-MODE=frame\_plus\_p2n.model -DEV=xl  
-Q/cgn2.1/USPTO\_spool/US09786713/runat.11082003.150513.6025/app\_query.fasta.1.1052  
-DB=N.Geneseq.19Jun03 -Qfmt=fastep -Suffix=ring -MINMATCH=0.1 -LOOPEL=0  
-LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blomsu62 -TRANS=human40.cdl  
-LIST=45 -DOCALLIGN=200 -THR\_SCORE=pct -THR\_MAX=100 -THR\_MIN=0 -ALIGN=15  
-MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000  
-USER=US09786713.ecgn1.1.874@runat.11082003.150513.6025 -NCPU=6 -ICPU=3  
-NO MAP -LARGEQUERY -NEG\_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG  
-DEV\_TIMEOUT=120 -WARN\_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6  
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :  
N.Geneseq.19Jun03:\*  
1: /SIDSI/gcgcdata/geneseq/geneseqn-emb1/NA1980.DAT:\*  
2: /SIDSI/gcgcdata/geneseq/geneseqn-emb1/NA1981.DAT:\*  
3: /SIDSI/gcgcdata/geneseq/geneseqn-emb1/NA1982.DAT:\*  
4: /SIDSI/gcgcdata/geneseq/geneseqn-emb1/NA1983.DAT:\*  
5: /SIDSI/gcgcdata/geneseq/geneseqn-emb1/NA1984.DAT:\*  
6: /SIDSI/gcgcdata/geneseq/geneseqn-emb1/NA1985.DAT:\*  
7: /SIDSI/gcgcdata/geneseq/geneseqn-emb1/NA1986.DAT:\*  
8: /SIDSI/gcgcdata/geneseq/geneseqn-emb1/NA1987.DAT:\*  
9: /SIDSI/gcgcdata/geneseq/geneseqn-emb1/NA1988.DAT:\*  
10: /SIDSI/gcgcdata/geneseq/geneseqn-emb1/NA1989.DAT:\*  
11: /SIDSI/gcgcdata/geneseq/geneseqn-emb1/NA1990.DAT:\*  
12: /SIDSI/gcgcdata/geneseq/geneseqn-emb1/NA1991.DAT:\*  
13: /SIDSI/gcgcdata/geneseq/geneseqn-emb1/NA1992.DAT:\*  
14: /SIDSI/gcgcdata/geneseq/geneseqn-emb1/NA1993.DAT:\*  
15: /SIDSI/gcgcdata/geneseq/geneseqn-emb1/NA1994.DAT:\*  
16: /SIDSI/gcgcdata/geneseq/geneseqn-emb1/NA1995.DAT:\*  
17: /SIDSI/gcgcdata/geneseq/geneseqn-emb1/NA1996.DAT:\*  
18: /SIDSI/gcgcdata/geneseq/geneseqn-emb1/NA1997.DAT:\*  
19: /SIDSI/gcgcdata/geneseq/geneseqn-emb1/NA1998.DAT:\*  
20: /SIDSI/gcgcdata/geneseq/geneseqn-emb1/NA1999.DAT:\*  
21: /SIDSI/gcgcdata/geneseq/geneseqn-emb1/NA2000.DAT:\*  
22: /SIDSI/gcgcdata/geneseq/geneseqn-emb1/NA2001A.DAT:\*  
23: /SIDSI/gcgcdata/geneseq/geneseqn-emb1/NA2001B.DAT:\*  
24: /SIDSI/gcgcdata/geneseq/geneseqn-emb1/NA2002.DAT:\*  
25: /SIDSI/gcgcdata/geneseq/geneseqn-emb1/NA2003.DAT:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,

and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Match	Length	DB	ID	Description
1	615	100.0	738	21	AAZ51739	Glycine max thior
2	485	78.9	3888	24	ABSS3097	DNA encoding Thior
3	485	78.9	3888	24	ABN89581	Phaseolin promoter
4	482	78.4	345	24	ABZ12358	Arabidopsis thalia
5	482	78.4	345	24	ABN89587	Arabidopsis thalia
6	482	78.4	561	21	AAC51522	Arabidopsis thalia
7	482	78.4	563	21	AAC34121	Arabidopsis thalia
8	482	78.4	3129	24	ABSS3095	DNA encoding thior
9	482	78.4	3129	24	ABN89579	Phaseolin promoter
10	482	78.4	3888	24	ABSS3096	DNA encoding Oleos
11	482	78.4	3888	24	ABN89580	Phaseolin promoter
12	478	77.7	470	21	AAC37781	Arabidopsis thalia
13	477	77.6	4935	24	ABN89586	Promoter-Oleosis t
14	472	76.7	509	22	AAH87768	Peppermint plant o
15	445	72.4	653	16	AAO99783	Plant SAR gene pl.
16	445	72.4	653	20	AAV62799	Tobacco SAR CHX in
17	445	72.4	653	20	AAV81683	Tobacco protein-sy
18	438	71.2	402	25	ABX21664	Human GDP-mannose
19	430	69.9	574	21	AAZ51738	Catalpa speciosa t
20	429	69.8	560	21	AAC41961	Arabidopsis thalia
21	424	68.9	392	25	ABX18045	Human GDP-mannose
22	410	66.7	601	21	AAZ51740	Glycine max thior
23	408	66.3	328	21	AAA31097	Plant microsateili
24	403	65.5	346	21	AAA31096	Plant microsateili
25	398	64.7	320	21	AAA31785	Plant microsateili
26	382	63.7	614	21	AAZ51741	Vernonia mespilifo
27	381	63.6	390	25	ABX19403	Human GDP-mannose
28	388	63.1	686	15	AAO78205	Gene coding for pr
29	388	63.1	687	21	AAC66375	Rice thioridoxin h
30	385.5	62.7	590	21	AAC63792	Arabidopsis thalia
31	381.5	62.0	360	24	ABZ12359	Arabidopsis thalia
32	376.5	61.2	357	24	ABZ13931	Arabidopsis thalia
33	376.5	61.2	480	21	AAC36542	Arabidopsis thalia
34	376.5	61.2	524	21	AAC33829	Arabidopsis thalia
35	376.5	61.2	563	21	AAC34211	Arabidopsis thalia
36	376.5	61.2	652	21	AAC48656	Arabidopsis thalia
37	376	61.1	292	21	AAA31978	Plant microsateili
38	368	59.8	393	17	AAT10451	Hard wheat thiorid
39	368	59.8	393	21	AAC62457	Wheat thioridoxin
40	367	59.7	382	21	AAC62456	Soft wheat thiorid
41	367	59.7	384	17	AAT10450	Plant microsateili
42	367	59.7	419	21	AAZ51942	Arabidopsis thalia
43	367	59.7	572	21	AAC52069	Arabidopsis thalia
44	366	59.5	870	21	AAZ51737	Momordica charanti
45	364.5	59.3	576	25	ABX56868	Arabidopsis thalia

## ALIGNMENTS

RESULT 1	AAZ51739
ID	AAZ51739 standard; cDNA: 738 BP.
AC	AAZ51739;
XX	XX
DT	04-JUL-2000 (first entry)
XX	XX
DE	Glycine max thioridoxin cDNA-1.
XX	XX
KW	Glycine max thioridoxin; clone sahlc-pk001.117; chimeric gene; soybean;
KW	transgenic plant; seed storage protein; allergenicity; ss.
XX	XX
OS	Glycine max.
XX	XX
FT	Key
CDS	Location/Qualifiers
	68..430

	FT	/tag- a	"Thioredoxin"
	FT	/product=	
XX	MW200014239-A2.		
XX	PN	16-MAR-2000.	
XX	PD		
XX	PF	07-SEP-1999; 99MO-US20420.	
XX	PR	08-SEP-1998; 98US-0099501.	
XX	(DUPLO ) DU PONT DE NEMOURS & CO E I.		
XX	Allan SM, Thorpe CJ, Lu AL;		
XX	WP1: 2000-256987/22.		
XX	P-PADB: AA170481.		
XX	New isolated polynucleotide encoding thioredoxin polypeptide is useful for producing transgenic plants with an altered level of thioredoxin -		
PS	Claim 3; Page 30; 33pp; English.		
CC	The present cDNA sequence encodes glycine max thioredoxin protein. The cDNA was derived from clone sahic.pk001.117, which was isolated from a cDNA library prepared from soybean tissue sprayed with authority herbicide. Chimeric genes encoding all or a portion of the thioredoxin protein, in sense or antisense orientation are constructed, wherein expression of the chimeric gene results in production of altered levels of the thioredoxin protein in a transformed host cell. Thioredoxin is involved in the disassembly of seed storage proteins during germination by reducing S-S bonds and in the bread making process. Over expression of thioredoxin in cereals may reduce the allergenicity of any transgenic protein engineered into cereal crops with high sulfhydryl content.		
SX	Sequence 738 BP; 260 A; 120 C; 146 G; 212 T; 0 other:		
Alignment Scores:			
Pred. No.:	4..7e-73	Length:	738
Score:	615.00	Matches:	120
Percent Similarity:	100.00%	Conservative:	0
Best local Similarity:	100.00%	Mismatches:	0
Query Match:	100.00%	Indels:	0
DB:	21	Gaps:	0
US-09-786-715-6 (1-120) x AAZ51739 (1-738)			
OY	1 MetAlAGISeSerScrlugLugInyValIIseSerCySHstHrVaIGluGlTPAsn 20		
Dd	68 ATGGCTGCGTCATCGAAGAAGGAACAAGCATTAAGTACCTCCACACC GTTGAAAGATTGAAC 127		
OY	21 AspGIuLeuGLInylsGLInuSlerLySLysleuIlleValJaSpheThrLaSer 40		
Dd	128 GATCAACTCAGAACGGCAAACAAATCCAAGAAACATGTGTTGGATTACTGCCTTCT 187		
OY	41 TrpCYvAgLIProCYsrArphelIlelabropheuJlaGluDeuaILalYSlySPhetR 60		
Dd	188 TGtGTGGkAccAtGCCgTTCATTCACCAATTCYTGGGTGAGCTGGCTAAAGACTTACA 247		
OY	61 SerVallIepheluLSyVALasPvalASpGiuleuLYSSerValISergINAsPTripALa 80		
Dd	248 AGTGCATATTCCTTAAGSTGTGATGTGCGCAATAAAGATGTTCTCAAGAATGGGCT 307		
OY	81 IllegualImetProthrPhryethaItheValylSGluGLYThrLeuEusApLySVaIVaI 100		
Dd	308 ATTGAAGCTATGCCCACTTTGTGTGTGTGAAGAAGGGAAAGCTTGGACAAAAGTGGTG 367		
OY	101.GLYAlaLyLSyASpGIuLeuGLInYLInySliegInLyShISvaIALaserIASerLA 120		
Dd	368 GGAGCAAAAGAGATGAGCTGCAGCAGAAAAATACGAAACATGTGGCTTAGCTAGTGTCT 427		
RESULT 2			
BBS53097			

ID		ABSS3097 standard; DNA: 3888 BP.
XX		
AC	ABSS3097;	
XX		
DT	29-NOV-2002	(first entry)
XX		
DE	DNA encoding Thioredoxin-oleosin fusion protein.	
XX		
KW	Thioredoxin; thioredoxin reductase; gene expression; oleosin; oil body; oleosin-thioredoxin fusion protein; gene; ds.	
XX		
OS	Arabidopsis thaliana.	
OS	Brassica napus.	
OS	Synthetic.	
XX		
PH	Key	Location/Qualifiers
FT	CDS	155..2658
FT	/tag= a	/product= "Thioredoxin-oleosin fusion protein"
FT	/product=	1555..2250
FT	exon	/tag= b
FT	/number= 1	2251..2489
FT	intron	/tag= c
FT	/number= 1	2490..2658
FT	exon	/tag= d
FT	/number= 2	
XX		
PM	US2002088025-A1.	
XX		
PD	04-JUL-2002.	
XX		
PF	03-JUL-2001;	2001US-0897425.
XX		
PR	22-FEB-1991;	91US-0659935.
PR	16-NOV-1993;	93US-0142418.
PR	30-DEC-1994;	94US-0366783.
PR	25-APR-1997;	97US-0846021.
PR	18-DEC-1998;	98US-0210843.
XX		
PA	(MOL/O) MOLONEY M.M. (DALM/) DALMIA B.K.	
P1	Moloney MM,	Dalmia BK;
X3		
DR	WPI; 2002-635723/68.	
XX		
PT	Expressing protein, by introducing chimeric nucleotide regulatory sequence, sequence encoding fusion protein, having sequence encoding protein, oleosin gene and sequence encoding termination region and producing protein -	
PS		
XX		
XX	Example 21; Fig 14; 69pp; English.	
CC	The invention describes a method of expressing thioredoxin or thioredoxin reductase (I) in the oil body of a host cell using an oil body protein gene. The method involves introducing a chimeric nucleic acid comprising a first sequence to regulate transcription, a second DNA sequence encoding a fusion polypeptide, comprising a sequence encoding an oleosin gene and sequence encoding (I), and a third sequence encoding a termination region functional in the host cell and growing the host cell to produce a fusion polypeptide. The method or (I) is useful for expression of a thioredoxin or thioredoxin reductase by a host cell. This sequence encodes a oleosin oil body protein fused to Arabidopsis thaliana thioredoxin gene controlled by a phaseolin promoter and phaseolin terminator sequence.	
CC		
CC		
CC		
CC		
SO	Sequence 3888 BP; 1254 A; 722 C; 613 G; 1299 T; 0 other;	
Alignment Scores:		
ered. No.:	1,58e-54	Length: 3888
score:	485.00	Matches: 89



Percent Similarity: 88.70% Conservative: 13  
 Best Local Similarity: 77.39% Mismatches: 13  
 Query Match: 78.86% Indels: 0  
 DB: 24 Gaps: 0

US-09-786-715-6 (1-120) x ABN89581 (1-3888)

QY 4 SerSerGluGluGluGluValIleSerCysHisThrValGluGluLutTrpAsnAspGlnLeu 23  
 DB 1558 GCTTCGGAAGAGACAGATGATCGCTGCCACACCGTTGAGACATGAGCAAGCAGAGCTT 1617  
 QY 24 GlnLysGlyAsnGluSerLysLysLeuIleValAlaValAspPheThrAlaSerTrpCysGly 43  
 DB 1618 CAGAAAGCTAATCAATCAAACTCTTGCTGGTGGTTCGATTCACGGCTTCTTGCTGCA 1677  
 QY 44 ProCysArgPheIleAlaProPheLeuAlaGluLeuAlaLysLysPheThrSerValIle 63  
 DB 1678 CCATGTCGTTTATCCCTCCATCTTGTGATTTGGCTAAGAACTTCTTACCTGCTT 1737  
 QY 64 PheLeuLysValAspValAspGluLeuLysSerValSerGlnAspTrpAlaIleGluAla 83  
 DB 1738 TTCCTCAAGTTGATCTGATGATGATGATGATGATGATGATGATGATGATGATGATG 1797  
 QY 84 MetProThrPheValPheValPheValLysGluGlyThrLeuLeuAspLysValAlaGlyAlaLys 103  
 DB 1798 ATGCCAACCTTCACTGTTTGAAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1857  
 QY 104 LysAspGluLeuGlnGlnLysIleGlnLysHisValAlaSerAla 118  
 DB 1858 AAGAGAGAGCTTCACTGATCTACATTCGCAAACTGCTGCTATGCGC 1902  
 RESULT 3  
 ABN89581  
 ID ABN89581 standard; DNA: 3888 BP.  
 XX  
 AC ABN89581;  
 XX  
 DT 06-SEP-2002 (first entry)  
 XX  
 DE Phaseolin promoter-Trrh oleosin-phaseolin terminator DNA SEQ:19.  
 XX  
 KM Multimeric protein; redox protein; thioredoxin; thioredoxin reductase;  
 KM oil body; ophthalmological; antidiabetic; cytoskeletal; antipsochalic;  
 KM vasotropic; vulnarary; antibacterial; immunosuppressive; anticancer;  
 KM food product; milk; wheat; oxidative stress; cataract; diabetes;  
 KM chronic obstructive pulmonary disease; envenomation; psoriasis; sepsis;  
 KM bronchiopulmonary disease; malignancy; reperfusion injury; wound healing;  
 KM gastro intestinal bleeding; intestinal bowel disease; ulcer;  
 KM gastro oesophageal reflux disease; gene; ds.  
 XX  
 OS Arabidopsis sp.  
 OS Phaseolus vulgaris.  
 XX  
 PN WO200250289-A1.  
 XX  
 PD 27-JUN-2002.  
 XX  
 PF 19-DEC-2001; 2001WO-US50240.  
 XX  
 PR 19-DEC-2000; 2000US-0742900.  
 PR 05-JUL-2001; 2001US-30285P.  
 PR 04-DEC-2001; 2001US-0006038.  
 XX  
 PA (SEMB-) SEMBIOSYS GENETICS INC.  
 PA (SYGN) SYNGENTA PARTICIPATIONS AG.  
 XX  
 PI Van Rooijjen G, Deckers H, Helfetz PB, Briggs SP, Dalmia BK;  
 PI Del Val G, Zaplachinski S, Moloney M.  
 XX  
 DR WPI, 2002-508806/54.  
 DR P-PSDB; ABP60683.  
 XX  
 PT Producing oil body associated with recombinant multimeric protein

PT complex e.g. redox proteins and immunoglobulins comprises producing  
 PT recombinant polypeptides capable of forming the complex in cells  
 PT comprising oil bodies -

XX Example 2: Page 169-171; 362pp; English.

XX The present invention describes a method (M1) for producing an oil body  
 CC associated with a recombinant multimeric protein complex (MPC). M1  
 CC comprises producing in a cell comprising oil bodies a first and second  
 CC recombinant polypeptide (P1, P2), where P1 is capable of associating  
 CC with P2 to form the MPC and associating the complex with an occlusion  
 CC body (OB) through an OB-targeting-protein capable of associating with OB  
 CC and P1. M1 is useful for producing an oil body associated with a  
 CC recombinant MPC. The oil bodies are further formulated for use in the  
 CC preparation of a food product such as milk or wheat based food product,  
 CC personal care product which reduces the oxidative stress on the surface  
 CC area of the human body or used to lighten the skin, or a pharmaceutical  
 CC composition used to treat chronic obstructive pulmonary disease (COPD),  
 CC cataracts, diabetes, envenomation, bronchiopulmonary disease, psoriasis,  
 CC malignancies, reperfusion injury, wound healing, sepsis, gastro  
 CC intestinal (GI) bleeding, intestinal bowel disease (IBD), ulcers, GERD  
 CC (gastro oesophageal reflux disease). ABN89581 to ABN89593 and ABP60677  
 CC to ABP60964 represent sequence given in the exemplification of the  
 CC present invention.  
 CC  
 XX

SO Sequence 3888 BP; 1254 A; 722 C; 613 G; 1299 T; 0 other;

#### Alignment Scores:

Pred. No.: 1,58e-54 Length: 3888  
 Score: 485.00 Matches: 89  
 Percent Similarity: 88.70% Conservative: 13  
 Best Local Similarity: 77.39% Mismatches: 13  
 Query Match: 78.86% Indels: 0  
 DB: 24 Gaps: 0

US-09-786-715-6 (1-120) x ABN89581 (1-3888)

QY 4 SerSerGluGluGluGluValIleSerCysHisThrValGluGluLutTrpAsnAspGlnLeu 23  
 DB 1558 GCTTCGGAAGAGACAGATGATCGCTGCCACACCGTTGAGACATGAGCAAGCAGAGCTT 1617  
 QY 24 GlnLysGlyAsnGluSerLysLysLeuIleValAlaValAspPheThrAlaSerTrpCysGly 43  
 DB 1618 CAGAAAGCTAATCAATCAAACTCTTGCTGGTGGTTCGATTCACGGCTTCTTGCTGCA 1677  
 QY 44 ProCysArgPheIleAlaProPheLeuAlaGluLeuAlaLysLysPheThrSerValIle 63  
 DB 1678 CCATGTCGTTTATCCCTCCATCTTGTGATTTGGCTAAGAACTTCTTACCTGCTT 1737  
 QY 64 PheLeuLysValAspValAspGluLeuLysSerValSerGlnAspTrpAlaIleGluAla 83  
 DB 1738 TTCCTCAAGTTGATCTGATGATGATGATGATGATGATGATGATGATGATGATGATG 1797  
 QY 84 MetProThrPheValPheValPheValLysGluGlyThrLeuLeuAspLysValAlaGlyAlaLys 103  
 DB 1798 ATGCCAACCTTCACTGTTTGAAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1857  
 QY 104 LysAspGluLeuGlnGlnLysIleGlnLysHisValAlaSerAla 118  
 DB 1858 AAGAGAGAGCTTCACTGATCTACATTCGCAAACTGCTGCTATGCGC 1902  
 RESULT 4  
 ABN89581  
 ID ABN89581 standard; DNA: 3888 BP.  
 XX  
 AC ABN89581;  
 XX  
 DT 21-JAN-2003 (first entry)  
 XX  
 DE Arabidopsis thaliana stress regulated gene SEQ ID NO 163.  
 XX  
 KM Arabidopsis thaliana; plant; gene; stress; transgenic; ds.  
 XX

OS Arabidopsis thaliana.  
 PN WO200216655-A2.  
 PD 28-FEB-2002.  
 XX 24-AUG-2001; 2001WO-US26685.  
 PF 24-AUG-2000; 2000US-227866P.  
 PR 26-JAN-2001; 2001US-264647P.  
 PR 22-JUN-2001; 2001US-300111P.  
 XX (SCRI ) SCRIPPS RES INST.  
 PA (SYGN ) SYNGENTA PARTICIPATIONS AG.  
 PI Harper JF, Kreps J, Wang X, Zhu T;  
 XX WPI: 2002-304127/34.  
 DR  
 XX Identifying a stress condition to which a plant cell has been exposed  
 PT and producing plants with increased tolerance to these abiotic stresses  
 PT -  
 XX  
 PS Claim 144: SEQ ID NO 163; 577pp + Sequence Listing: English.  
 CC The invention relates to identifying a stress condition to which a plant  
 CC cell has been exposed, comprising:  
 CC (a) contacting nucleic acid representative of expressed polynucleotides  
 CC in the plant cell with an array or probes representative of the plant  
 CC cell genome; and  
 CC (b) detecting a profile of expressed polynucleotides in the plant cell  
 CC characteristic of a stress response. The method is useful in the  
 CC production of transgenic plants, cells and seeds and in producing plants  
 CC with increased tolerance to abiotic stress. The present sequence is that  
 CC of an Arabidopsis thaliana stress regulated gene (AB212196-AB21574) used  
 CC in methods of the invention.  
 CC Note: The sequence data for this patent is not represented in the printed  
 CC specification but is based on sequence information supplied to Derwent by  
 CC the European Patent Office.  
 CC  
 XX  
 SQ Sequence 345 BP; 90 A; 68 C; 90 G; 97 T; 0 other;  
 Alignment Scores:  
 Pred. No.: 1,32e-55 Length: 345  
 Score: 482.00 Matches: 88  
 Percent Similarity: 89.388 Conservative: 13  
 Best Local Similarity: 77.888 Mismatches: 12  
 Query Match: 78.376 Indels: 0  
 DB: 24 Gaps: 0  
 US-09-786-715-6 (1-120) x AB212358 (1-345)  
 QY 4 SerSerGluGluGluGluValIleSerCysHisThrValGluGluTyrPAsnAspGlnLeu 23  
 Db 4 GCTTCGGAAGAAAGACAAAGTGAATCCGCGCCACACCGTTGAAGATGGAAAGACACACTT 63  
 QY 24 GlnLysGlyValAsnGluSerLysLysLeuIleValAlaSerPheThrAlaSerTyrCysGly 43  
 Db 64 CAAAGAGCATGAATCAACCAAACTCTTGCGGTGATTTACAGCGCTCTGSGTGGCA 123  
 QY 44 ProCysArgPheIleAlaProPheLeuAlaGluLeuAlaLysLysPheThrSerValIle 63  
 Db 124 CCATGTCGTTTATCCCTCATTTCTTCTGATTTGGCTAGCAAACTCCATCAAGTCTT 183  
 QY 64 PheLeuLysValAlaSpluAlaSpluLeuLysSerValSerGlnAspTTPAlaIleGluAla 83  
 Db 184 TTCCTCAAGGTTGATACGTGATGAATGAAGTGGCGACAGTGAATGGCGATACAGCGC 243  
 QY 84 MetProThrPheValPheValLysGluGlyThrLeuLeuAspLysValAlaGlyAlaLys 103  
 Db 244 ATGCCAACCTCATATGTTTTGAAGGAAGGAAGATTTTGGACAAAGTTGTTGGACCAAG 303  
 QY 104 LysAspGluLeuGlnGlnLysIleGlnLysHisValAla 116

DB	304	AAAGATGACCTTCAGCTTACCACTTGGCAAAACACTTGGCT	342
RESULT 5			
ABN89587			
ID	ABN89587	standard; DNA: 345 BP.	
XX			
AC	ABN89587;		
XX			
DT	06-SEP-2002	(first entry)	
XX			
DE	Arabidopsis thaliana thioresoxin h (Trx h 1) DNA SEQ ID NO:38.		
KW	Multimeric protein; redox protein; thioresoxin; thioresoxin reductase;		
KW	oil body; ophthalmological; antidiabetic; cytosol; antiproliferative;		
KW	vasotrophic; vulnery; antibacterial; immunosuppressive; antitumor;		
KW	food product; milk; wheat; oxidative stress; cataract; diabetes;		
KW	chronic obstructive pulmonary disease; emphysema; psoriasis; sepsis;		
KW	bronchiopulmonary disease; malignancy; reperfusion injury; wound healing;		
KW	gastro intestinal bleeding; intestinal bowel disease; ulcer;		
KW	gastro oesophageal reflux disease; gene; ds.		
XX			
OS	Arabidopsis thaliana.		
XX			
PN	WO200250289-A1.		
XX			
PD	27-JUN-2002.		
XX			
PF	19-DEC-2001; 2001WO-US050240.		
XX			
FR	19-DEC-2000; 2000US-0742800.		
PR	05-JUL-2001; 2001US-302885P.		
XX			
PR	04-DEC-2001; 2001US-0006038.		
XX			
PA	(SEMB-) SEMBIOSYS GENETICS INC.		
XX	(SYGN ) SYNGENTA PARTICIPATIONS AG.		
PI	Van Rooijen G, Deckers H, Heifetz PB, Briggs SP, Dalmia BK;		
PI	Del Val G, Zaplachinski S, Moloney M;		
DR	WPI: 2002-508806/54.		
XX	P-PSDB; ABP60696.		
PT	Producing oil body associated with recombinant multimeric protein		
PT	complex e.g. redox proteins and immunoglobulins comprises producing		
PT	recombinant polypeptides capable of forming the complex in cells		
PT	comprising oil bodies -		
XX			
PS	Claim 68; Page 197; 362pp; English.		
XX			
CC	The present invention describes a method (M1) for producing an oil body		
CC	associated with a recombinant multimeric protein complex (MPC). M1		
CC	comprises producing in a cell comprising oil bodies a first and second		
CC	recombinant polypeptide (P1, P2), where P1 is capable of associating		
CC	with P2 to form the MPC and associating the complex with an occlusion		
CC	body (OB) through an OB-targeting-protein capable of associating with OB		
CC	and P1. M1 is useful for producing an oil body associated with a		
CC	recombinant MPC. The oil bodies are further formulated for use in the		
CC	preparation of a food product such as milk or wheat based food product,		
CC	personal care product which reduces the oxidative stress on the surface		
CC	area of the human body or used to lighten the skin, or a pharmaceutical		
CC	composition used to treat chronic obstructive-pulmonary disease (COPD),		
CC	cataracts, diabetes, emphysema, bronchiopulmonary disease, psoriasis,		
CC	malinancies, reperfusion injury, wound healing, sepsis, gastro		
CC	intestinal (GI) bleeding, intestinal bowel disease (IBD), ulcers, GEND		
CC	(gastro oesophageal reflux disease). ABN89589 to ABN89593 and ABP60677		
CC	to ABP60964 represent sequence given in the exemplification of the		
CC	present invention.		
XX			
SO	Sequence 345 BP; 90 A; 68 C; 90 G; 97 T; 0 other;		



```

PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 09-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149375.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 25-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151085.
PR 27-AUG-1999; 99US-0151086.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159283.
PR 13-OCT-1999; 99US-0159284.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.

```

```

PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

```

```

Alignment Scores:
Pred. No.: 2,62e-55 Length: 561
Score: 482.00 Matches: 88
Percent Similarity: 89.38% Conservative: 13
Best Local Similarity: 77.88% Mismatches: 12
Query Match: 78.37% Indels: 0
DB: 21 Gaps: 0

```

US-09-786-715-6 (1-120) x AAC51522 (1-561)

```

QY 4 SerserGlucIuGlyGlnValIleserCyHisThrValGluGluTrpAsnAspGlnLeu 23
DB 97 GCTTGGAGAGACAGACAGATGATCGCTGCCACACCTTGAACATGAGATGAGACAGACTT 156
QY 24 GlnIysGlyAsnGlySerLysLysLeuIleValValAspPheThrAlaSerTPCyGly 43
DB 157 CAGAGGCTATATATATCAAAACTCTGTGTGTGATTTCACGCGCTTCTGTGTGGA 216
QY 44 ProCysArpHeIleAlaProPheLeuAlaGlnLeuAlaLysLysPheThrIleValIle 63
DB 217 CCATGTGTTTCATCGCTCCATCTTGTGATTTGGCTAAGAACTTCTTAAGCTGCTT 276
QY 64 PheLeuLysValAspValAspGluLeuLysSerValSerGlnAspTrpAlaIleGluAla 83
DB 277 TTCCTCAAGCTTGAATGATGATGAATGATGATGATGATGATGATGATGATGATGATGATG 336
QY 84 MetProThrPheValPheValLysGluGlyThrLeuLeuAspLysValValIleGlyAlaLys 103
DB 337 ATGCCAACCTTTCATGTTTGAAGAGAGCAATTTTGGACAAGTTTGGAGCCAG 396
QY 104 LysAspGluLeuGlnLysIleGlnLysIleValAla 116
DB 397 AAGATGAGCTTCATGCTACCATTTGCCAACAACCTTGCT 435
RESULT 7
AAC34121
ID AAC34121. standard; DNA; 563 BP.
XX
AC AAC34121.
XX
XX 17-OCT-2000 (first entry)
XX
DE Arabidopsis thaliana DNA fragment SEQ ID NO: 5524.
XX
XX Hybridisation assay; genetic mapping; gene expression control;
KW protein identification; signal transduction pathway;
KW metabolic pathway; promoter; termination sequence; ss.
XX
OS Arabidopsis thaliana.
XX
PN EP1033405-A2.
XX
PD 06-SEP-2000.
XX
PF 25-FEB-2000; 2000EP-0301439.
XX
PR 25-FEB-1999; 99US-0121825.

```

PR 05-MAR-1999; 99US-0123180.  
PR 09-MAR-1999; 99US-0123548.  
PR 23-MAR-1999; 99US-0125788.  
PR 25-MAR-1999; 99US-0126264.  
PR 29-MAR-1999; 99US-0126785.  
PR 01-APR-1999; 99US-0127462.  
PR 06-APR-1999; 99US-0128234.  
PR 08-APR-1999; 99US-0128714.  
PR 16-APR-1999; 99US-0129845.  
PR 19-APR-1999; 99US-0130077.  
PR 21-APR-1999; 99US-0130449.  
PR 23-APR-1999; 99US-0130510.  
PR 28-APR-1999; 99US-0130891.  
PR 30-APR-1999; 99US-0131449.  
PR 30-APR-1999; 99US-0132048.  
PR 04-MAY-1999; 99US-0132407.  
PR 05-MAY-1999; 99US-0132484.  
PR 06-MAY-1999; 99US-0132485.  
PR 06-MAY-1999; 99US-0132486.  
PR 07-MAY-1999; 99US-0132487.  
PR 11-MAY-1999; 99US-0132863.  
PR 14-MAY-1999; 99US-0134256.  
PR 14-MAY-1999; 99US-0134218.  
PR 14-MAY-1999; 99US-0134219.  
PR 14-MAY-1999; 99US-0134421.  
PR 18-MAY-1999; 99US-0134370.  
PR 18-MAY-1999; 99US-0134768.  
PR 19-MAY-1999; 99US-0134941.  
PR 20-MAY-1999; 99US-0135124.  
PR 21-MAY-1999; 99US-0135353.  
PR 24-MAY-1999; 99US-0135629.  
PR 25-MAY-1999; 99US-0136021.  
PR 27-MAY-1999; 99US-0136392.  
PR 28-MAY-1999; 99US-0136782.  
PR 01-JUN-1999; 99US-0137222.  
PR 03-JUN-1999; 99US-0137528.  
PR 04-JUN-1999; 99US-0137502.  
PR 07-JUN-1999; 99US-0137724.  
PR 08-JUN-1999; 99US-0138094.  
PR 10-JUN-1999; 99US-0138540.  
PR 10-JUN-1999; 99US-0138847.  
PR 14-JUN-1999; 99US-0139119.  
PR 16-JUN-1999; 99US-0139452.  
PR 17-JUN-1999; 99US-0139453.  
PR 17-JUN-1999; 99US-0139492.  
PR 18-JUN-1999; 99US-0139454.  
PR 18-JUN-1999; 99US-0139455.  
PR 18-JUN-1999; 99US-0139456.  
PR 18-JUN-1999; 99US-0139457.  
PR 18-JUN-1999; 99US-0139458.  
PR 18-JUN-1999; 99US-0139459.  
PR 18-JUN-1999; 99US-0139460.  
PR 18-JUN-1999; 99US-0139461.  
PR 18-JUN-1999; 99US-0139462.  
PR 18-JUN-1999; 99US-0139463.  
PR 18-JUN-1999; 99US-0139750.  
PR 18-JUN-1999; 99US-0139751.  
PR 21-JUN-1999; 99US-0139817.  
PR 22-JUN-1999; 99US-0139899.  
PR 23-JUN-1999; 99US-0140353.  
PR 23-JUN-1999; 99US-0140354.  
PR 24-JUN-1999; 99US-0140695.  
PR 28-JUN-1999; 99US-0140823.  
PR 29-JUN-1999; 99US-0140991.  
PR 30-JUN-1999; 99US-0141827.  
PR 01-JUL-1999; 99US-0141842.  
PR 02-JUL-1999; 99US-0142154.  
PR 06-JUL-1999; 99US-0142055.  
PR 08-JUL-1999; 99US-0142390.  
PR 09-JUL-1999; 99US-0142803.  
PR 09-JUL-1999; 99US-0142920.  
PR 12-JUL-1999; 99US-0142977.  
PR 13-JUL-1999; 99US-0143542.

PR 14-JUL-1999; 99US-0143624.  
PR 15-JUL-1999; 99US-0144005.  
PR 16-JUL-1999; 99US-0144085.  
PR 16-JUL-1999; 99US-0144086.  
PR 19-JUL-1999; 99US-0144325.  
PR 19-JUL-1999; 99US-0144331.  
PR 19-JUL-1999; 99US-0144332.  
PR 19-JUL-1999; 99US-0144333.  
PR 19-JUL-1999; 99US-0144334.  
PR 19-JUL-1999; 99US-0144335.  
PR 20-JUL-1999; 99US-0144352.  
PR 20-JUL-1999; 99US-0144632.  
PR 20-JUL-1999; 99US-0144884.  
PR 21-JUL-1999; 99US-0144814.  
PR 21-JUL-1999; 99US-0145086.  
PR 21-JUL-1999; 99US-0145088.  
PR 22-JUL-1999; 99US-0145085.  
PR 22-JUL-1999; 99US-0145087.  
PR 22-JUL-1999; 99US-0145089.  
PR 22-JUL-1999; 99US-0145192.  
PR 23-JUL-1999; 99US-0145145.  
PR 23-JUL-1999; 99US-0145218.  
PR 26-JUL-1999; 99US-0145224.  
PR 27-JUL-1999; 99US-0145276.  
PR 27-JUL-1999; 99US-0145913.  
PR 27-JUL-1999; 99US-0145918.  
PR 27-JUL-1999; 99US-0145919.  
PR 28-JUL-1999; 99US-0145951.  
PR 02-AUG-1999; 99US-0146386.  
PR 02-AUG-1999; 99US-0146388.  
PR 02-AUG-1999; 99US-0146389.  
PR 03-AUG-1999; 99US-0147038.  
PR 04-AUG-1999; 99US-0147204.  
PR 04-AUG-1999; 99US-0147302.  
PR 05-AUG-1999; 99US-0147192.  
PR 05-AUG-1999; 99US-0147260.  
PR 06-AUG-1999; 99US-0147303.  
PR 06-AUG-1999; 99US-0147416.  
PR 09-AUG-1999; 99US-0147493.  
PR 09-AUG-1999; 99US-0147935.  
PR 10-AUG-1999; 99US-0148171.  
PR 11-AUG-1999; 99US-0148319.  
PR 12-AUG-1999; 99US-0148341.  
PR 13-AUG-1999; 99US-0148565.  
PR 13-AUG-1999; 99US-0148684.  
PR 16-AUG-1999; 99US-0149368.  
PR 17-AUG-1999; 99US-0149175.  
PR 18-AUG-1999; 99US-0149426.  
PR 20-AUG-1999; 99US-0149722.  
PR 20-AUG-1999; 99US-0149723.  
PR 20-AUG-1999; 99US-0149929.  
PR 23-AUG-1999; 99US-0149902.  
PR 23-AUG-1999; 99US-0149930.  
PR 25-AUG-1999; 99US-0150566.  
PR 26-AUG-1999; 99US-0150884.  
PR 27-AUG-1999; 99US-0151065.  
PR 27-AUG-1999; 99US-0151066.  
PR 27-AUG-1999; 99US-0151080.  
PR 30-AUG-1999; 99US-0151303.  
PR 31-AUG-1999; 99US-0151438.  
PR 01-SEP-1999; 99US-0151930.  
PR 07-SEP-1999; 99US-0152363.  
PR 10-SEP-1999; 99US-0153070.  
PR 13-SEP-1999; 99US-0153758.  
PR 15-SEP-1999; 99US-0154018.  
PR 16-SEP-1999; 99US-0154039.  
PR 20-SEP-1999; 99US-0154779.  
PR 22-SEP-1999; 99US-0155139.  
PR 23-SEP-1999; 99US-0155486.  
PR 24-SEP-1999; 99US-0155659.  
PR 28-SEP-1999; 99US-0156458.  
PR 29-SEP-1999; 99US-0156596.  
PR 04-OCT-1999; 99US-0157117.

PR 05-OCT-1999; 99US-0157753.  
PR 06-OCT-1999; 99US-0157865.  
PR 07-OCT-1999; 99US-0158029.  
PR 08-OCT-1999; 99US-0158232.  
PR 12-OCT-1999; 99US-0158369.  
PR 13-OCT-1999; 99US-0159293.  
PR 13-OCT-1999; 99US-0159294.  
PR 13-OCT-1999; 99US-0159295.  
PR 14-OCT-1999; 99US-0159329.  
PR 14-OCT-1999; 99US-0159330.  
PR 14-OCT-1999; 99US-0159331.  
PR 14-OCT-1999; 99US-0159637.  
PR 14-OCT-1999; 99US-0159638.  
PR 18-OCT-1999; 99US-0159584.  
PR 21-OCT-1999; 99US-0160741.  
PR 21-OCT-1999; 99US-0160767.  
PR 21-OCT-1999; 99US-0160770.  
PR 21-OCT-1999; 99US-0160814.  
PR 21-OCT-1999; 99US-0160815.  
PR 22-OCT-1999; 99US-0160980.  
PR 22-OCT-1999; 99US-0160981.  
PR 22-OCT-1999; 99US-0160989.  
PR 25-OCT-1999; 99US-0161404.  
PR 25-OCT-1999; 99US-0161405.  
PR 25-OCT-1999; 99US-0161406.  
PR 26-OCT-1999; 99US-0161359.  
PR 26-OCT-1999; 99US-0161360.  
PR 26-OCT-1999; 99US-0161361.  
PR 28-OCT-1999; 99US-0161920.  
PR 28-OCT-1999; 99US-0161992.  
PR 28-OCT-1999; 99US-0161993.  
PR 29-OCT-1999; 99US-0162142.

Alignment Scores:  
Pred. No.: 2.63e-55 Length: 563  
Score: 482.00 Matches: 88  
Percent Similarity: 89.38% Conservative: 13  
Best Local Similarity: 77.88% Mismatches: 12  
Query Match: 78.37% Indels: 0  
DB: 21 Gaps: 0

US-09-786-715-6 (1-120) x AAC34121 (1-563)

OY 4 SerSerGluGluGluGlnValIleSerCySHsThrValGluGluTrpAsnAspGlnLeu 23  
DB 99 GCTTCGGAAGAGACAGATGATGCGCTGCCACACCGTTGAGACATGGAGACGACGCTT 158  
OY 24 GlnIysGlyAsnGluSerLysLysLeuIleValAlaAspPheThrAlaSerTrpCysGly 43  
DB 159 CAGAAAGGCTATGATCCAAACCTCTTGTTGGTGTATTCACGGCTTCTTGGGTGGA 218  
OY 44 ProCysArgPheIleAlaProPheLeuAlaGluLeuAlaLysLysPheThServalIle 63  
DB 219 CCAATGCTGTTTCATCGCTCCATCTTCTGATTTGGCTAAGAAACCTCCAAACGTCCTT 278  
OY 64 PheIleuysValaAspValaAspGluLeuysSerValSerGlnAspTrpAlaIleGluAla 83  
DB 279 TTCTCTCAAGGTGTAAGTGAATGATGAAGTGGTGGCAAGTATGGCGATACAGGCG 338  
OY 84 MetProThrPheValPheValIysGluGlyThrLeuLeuAspLysValaIleAlaLys 103  
DB 339 ATGCCAACCTTCACTTTTGAAGAGGAAGATTTTGGACAAAGTTGTGGAGCCAG 398  
OY 104 LysAspGluLeuGlnGlnLysIleGlnLysHisValAla 116  
DB 399 AAGATGAGCTTCAGTCTACCATTTGCCAAACACTTGGCT 437

RESULT 8  
ABSS3095  
ID ABSS3095 standard; DNA: 3129 BP.  
XX  
AC ABSS3095;

XX 29-NOV-2002 (first entry)  
DT DNA encoding thioredoxin-phaseolin fusion protein.  
DE  
XX Thioredoxin; thioredoxin reductase; gene expression; oleosin;  
XX oil body; gene; ds; thioredoxin reductase; phaseolin.  
KW  
XX Arabidopsis thaliana.  
OS Phaseolin vulgaris.  
OS Synthetic.  
XX  
XX Key Location/Qualifiers  
FH 1555..1899  
FT CDS /\*tag=a  
FT /product="thioredoxin-phaseolin fusion protein"

US200208025-A1.

04-JUL-2002.

03-JUL-2001; 2001US-0897425.

XX 22-FEB-1991; 91US-0659835.  
PR 16-NOV-1993; 93US-0142418.  
PR 30-DEC-1994; 94US-0366783.  
PR 25-APR-1997; 97US-0846021.  
PR 18-DEC-1998; 98US-0210843.

PA (MOLO/) MOLONEY M. M.  
PA (DALM/) DALMIA B. K.

XX Moloney MM, Dalmla BK;

DR WPI: 2002-635723/68.  
XX P-PSDB: ABG32916.

PT Expressing protein, by introducing chimeric nucleotide regulatory  
PT sequence, sequence encoding fusion protein, having sequence encoding  
PT protein, oleosin gene and sequence encoding termination region and  
PT producing protein -

XX Example 21; Fig 12; 69pp; English.

XX The invention describes a method of expressing thioredoxin or thioredoxin  
CC reductase (I) in the oil body of a host cell using an oil body protein  
CC gene. The method involves introducing a chimeric nucleic acid comprising  
CC a first sequence to regulate transcription, a second DNA sequence  
CC encoding a fusion polypeptide, comprising a sequence encoding an oleosin  
CC gene and sequence encoding (I) and a third sequence encoding a  
CC termination region functional in the host cell and growing the host cell  
CC to produce a fusion polypeptide. The method or (I) is useful for  
CC expression of a thioredoxin or thioredoxin reductase by a host cell. This  
CC sequence represents an Arabidopsis thaliana thioredoxin gene controlled  
CC by a phaseolin promoter and terminator for use in an expression vector.

XX Sequence 3129 BP; 1051 A; 549 C; 439 G; 1090 T; 0 other;

Alignment Scores:  
Pred. No.: 2.95e-54 Length: 3129  
Score: 482.00 Matches: 88  
Percent Similarity: 89.38% Conservative: 13  
Best Local Similarity: 77.88% Mismatches: 12  
Query Match: 78.37% Indels: 0  
DB: 24 Gaps: 0

US-09-786-715-6 (1-120) x ABSS3095 (1-3129)

OY 4 SerSerGluGluGluGlnValIleSerCySHsThrValGluGluTrpAsnAspGlnLeu 23  
DB 1558 GCTTCGGAAGAGACAGATGATGCGCTGCCACACCGTTGAGACATGGAGACGACGCTT 1617  
OY 24 GlnIysGlyAsnGluSerLysLysLeuIleValAlaAspPheThrAlaSerTrpCysGly 43

```

DB      1618  CAGAAAGCTAATGAATCAAAACCTTGTGGTGTGATTTACGCGCTTCTGTGTGGA 1677
QY      44  ProCysArgPheIleAlaProPheLeuAlaIleAlaLysLysPheThrSerValIle 63
DB      1678  CCATGCGCTTTATCCCTCCATCTTGTGATTGGCTAAGAAACCTTCTACCGTGCCT 1737
QY      64  PheLeuValAspValAspValAspLeuLysSerValSerGlnAspTrpAlaIleGluAla 83
DB      1738  TTCCTCAAGCTTGATCTGATGATGATTAAGTCCGCTGGCAAGTATTGGCGGATACAGGCC 1797
QY      84  MetProThrPheValPheValLysGluGlyThrLeuLeuAspLysValValGlyAlaLys 103
DB      1798  ATGCCAACCTTCATGTTTGTGAAGGAAGCAAGATTTTGACAAAGTTGTGGAGCCCAAG 1857
QY      104  LysAspGluLeuGlnGlnIleLysIleGlnLysHisValAla 116
DB      1858  AAGATGAGCTTCAGTCTACCATTTGCCAAACACTTGGCT 1896
RESULT 9
ABN89579 standard; DNA; 3129 BP.
AC      ABN89579;
XX      06-SEP-2002 (first entry)
XX      Phascolin promoter-Arabidopsis Trxn-phascolin terminator DNA SEQ.14.
DE      Multimeric protein; redox protein; thioredoxin; thioredoxin reductase;
XX      oil body; ophthalmological; antidiabetic; cytoskeletal; antipsoriatic;
XX      vasotropic; vulnary; antibacterial; immunosuppressive; antilucer;
XX      food product; milk; wheat; oxidative stress; cataract; diabetes;
XX      chronic obstructive pulmonary disease; envenomation; psoriasis; sepsis;
XX      bronchiopulmonary disease; malignancy; reperfusion injury; wound healing;
XX      gastro intestinal bleeding; intestinal bowel disease; ulcer;
XX      gastro oesophageal reflux disease; gene; ds.
OS      Arabidopsis sp.
OS      Phascolus vulgaris.
PN      WO200250289-A1.
XX      27-JUN-2002.
XX      19-DEC-2001: 2001MO-US50240.
PF      19-DEC-2000: 2000US-0742900.
PR      05-JUL-2001: 2001US-302885P.
PR      04-DEC-2001: 2001US-0006038.
XX      (SEMB-) SEMBIOSYS GENETICS INC.
PA      (SYGN) SYNGENTA PARTICIPATIONS AG.
XX      Van Rooijen G, Deckers H, Helfetz PB, Briggs SP, Dalmia BK;
PI      Del Val G, Zaplachinski S, Moloney M;
XX      MPI: 2002-508806/54.
DR      P-PSDB; ABP60680.
XX      Producing oil body associated with recombinant multimeric protein
XX      complex e.g. redox proteins and immunoglobulins comprises producing
XX      recombinant polypeptides capable of forming the complex in cells
XX      comprising oil bodies -
PS      Example 2; Page 165-166; 362pp; English.
XX      The present invention describes a method (M1) for producing an oil body
XX      associated with a recombinant multimeric protein complex (MPC). M1
XX      comprises producing in a cell comprising oil bodies a first and second
XX      recombinant polypeptide (P1, P2), where P1 is capable of associating
XX      with P2 to form the MPC and associating the complex with an occlusion
XX      body (OB) through an OB-targeting-protein capable of associating with OB

```

```

CC      and P1. M1 is useful for producing an oil body associated with a
CC      recombinant MPC. The oil bodies are further formulated for use in the
CC      preparation of a food product such as milk or wheat based food product,
CC      personal care product which reduces the oxidative stress on the surface
CC      area of the human body or used to lighten the skin, or a pharmaceutical
CC      composition used to treat chronic obstructive pulmonary disease (COPD),
CC      cataracts, diabetes, envenomation, bronchiopulmonary disease, psoriasis,
CC      malignancies, reperfusion injury, wound healing, sepsis, gastro
CC      intestinal (GI) bleeding, intestinal bowel disease (IBD), ulcers, GERD
CC      (gastro oesophageal reflux disease). ABN89569 to ABN89593 and ABP60677
CC      to ABP60964 represent sequence given in the exemplification of the
CC      present invention.
SQ      Sequence 3129 BP; 1051 A; 549 C; 439 G; 1090 T; 0 other;

Alignment Scores:
Pred. No.: 2.95e-54 Length: 3129
Score: 482.00 Matches: 88
Percent Similarity: 89.38% Conservative: 13
Best Local Similarity: 77.88% Mismatches: 12
Query Match: 78.37% Indels: 0
DB: 24 Gaps: 0

US-09-786-715-6 (1-120) x ABN89579 (1-3129)
QY      4  SerSerGluGluGlnGlnValIleSerCysHisThrValGluGluTrpAsnAspGlnLeu 23
DB      1558  GCTTCGGAGACAGAGACAGATGATCGCTGCCACACCGCTTGACGATGAGACAGCAGCTT 1617
QY      24  GlnLysGlyAsnGluSerLysLysLeuIleValValAspPheThrAlaSerTrpCysGly 43
DB      1618  CAGAAAGCTAATGAATCAAAACCTTGTGGTGTGATTTACGCGCTTCTGTGTGGA 1677
QY      44  ProCysArgPheIleAlaProPheLeuAlaIleAlaLysLysPheThrSerValIle 63
DB      1678  CCATGCGCTTTATCCCTCCATCTTGTGATTGGCTAAGAAACCTTCTACCGTGCCT 1737
QY      64  PheLeuValAspValAspValAspLeuLysSerValSerGlnAspTrpAlaIleGluAla 83
DB      1738  TTCCTCAAGCTTGATCTGATGATGATTAAGTCCGCTGGCAAGTATTGGCGGATACAGGCC 1797
QY      84  MetProThrPheValPheValLysGluGlyThrLeuLeuAspLysValValGlyAlaLys 103
DB      1798  ATGCCAACCTTCATGTTTGTGAAGGAAGCAAGATTTTGACAAAGTTGTGGAGCCCAAG 1857
QY      104  LysAspGluLeuGlnGlnIleLysIleGlnLysHisValAla 116
DB      1858  AAGATGAGCTTCAGTCTACCATTTGCCAAACACTTGGCT 1896
RESULT 10
ABS53096
ID      ABS53096 standard; DNA; 3888 BP.
XX      ABS53096;
AC      29-NOV-2002 (first entry)
XX      DNA encoding Oleosin-thioredoxin fusion protein.
DE      Thioredoxin; thioredoxin reductase; gene expression; oleosin;
XX      oil body; oleosin-thioredoxin fusion protein; gene; ds.
XX      Arabidopsis thaliana.
OS      Brassica napus.
OS      Synthetic.
XX      Key Location/Qualifiers
FH      CDS 1555..2658
FT      /*tag= a
FT      /product= "Oleosin-thioredoxin fusion protein"
FT      1555..1907
FT      /*tag= b
FT      /number= 1

```

```

FT      Intron      1908..2147
FT      /*tag= c
FT      /number= 1
FT      exon      2148..2658
FT      /*tag= d
FT      /number= 2
XX      US2002088025-A1.
XX      04-JUL-2002.
XX      03-JUL-2001; 2001US-0897425.
XX      22-FEB-1991; 91US-0659835.
XX      16-NOV-1993; 93US-0142418.
XX      30-DEC-1994; 94US-0366783.
XX      25-APR-1997; 97US-0846021.
XX      18-DEC-1998; 98US-0210843.
XX      (MOLO/) MOLONEY M. M.
XX      (DALM/) DALMIA B. K.
XX      Moloney KM, Dalmia BK;
XX      WPI: 2002-635723/68.
XX      P-PSDB: AB632917.
XX      Expressing protein, by introducing chimeric nucleotide regulatory
XX      protein, sequence encoding fusion protein, having sequence encoding
XX      protein, oleosin gene and sequence encoding termination region and
XX      producing protein.
XX      Example 21: Fig 13; 69pp; English.
XX      The invention describes a method of expressing thioredoxin or thioredoxin
XX      reductase (I) in the oil body of a host cell using an oil body protein
XX      gene. The method involves introducing a chimeric nucleic acid comprising
XX      a first sequence to regulate transcription, a second DNA sequence
XX      encoding a fusion polypeptide, comprising a sequence encoding an oleosin
XX      gene and sequence encoding (I) and a third sequence encoding a
XX      termination region functional in the host cell and growing the host cell
XX      to produce a fusion polypeptide. The method or (I) is useful for
XX      expression of a thioredoxin or thioredoxin reductase by a host cell. This
XX      sequence encodes a oleosin oil body protein fused to Arabidopsis thaliana
XX      thioredoxin gene controlled by a phaseolin promoter and phaseolin
XX      terminator sequence.
XX      Sequence 3888 BP; 1254 A; 723 C; 613 G; 1298 T; 0 other;
SQ      Alignment Scores:
Pred. No.: 4,01e-54 Length: 3888
Score: 482.00 Matches: 88
Percent Similarity: 89.38% Conservative: 13
Best Local Similarity: 77.88% Mismatches: 12
Query Match: 78.37% Indels: 0
DB: 24 Gaps: 0
US-09-786-715-6 (1-120) x ABS53096 (1-3888)
QY      4 SererGIUGLUGLUGLValIIleSerCYSHSThrValIGLUTrpAsnAspGluLeu 23
DB      2317 GCTTCGGAAGAGGACAGTATCGCCGACACACCGTTGAGACATGGAGACGAGAGCTT 2376
QY      24 GlnrSGIyASngLserLySLysLeuIIleValIAspPhetHrAlaSerTrpCysGly 43
DB      2377 CAGAAAGCTATGATCAATCAAAACCTTGTGTGTGTTTTCACGCGCTTCTTGCTGTA 2436
QY      44 ProCysArgPheIIleAlaProPheLeuAlaGluLeuAlaLysLysPheHrSerValIle 63
DB      2437 CCATGTCGTTTCATTCGTCATTCCTTCTGATTTGGCTAGAAACCTTCTTAAGCGTCTT 2496
QY      64 PheLeuLysValaLysPheValaLysPheLysSerValaLysPheLysPheLysPheLys 83

```

```

DB      2497 TTCCTCAAGCTTGATGATGATGAATGAATGCGTGGCAAGTATGGCGCATACAGCGC 2556
QY      84 MetProThrPheValPheValLysGluGlyThrLeuLeuAspLysValaLysValaLys 103
DB      2557 ATGCCACCTTCATGTGTTTGGAGGAGGAGATTTTGGACAAAGTTGTTGGAGCCAAAG 2616
QY      104 LysAspGluLeuGlnGlnLysIleGlnLysHsValaIa 116
DB      2617 AAGATGAGCTTCAGTCTACATTCGCAACACTGGCGT 2655
RESULT 11
ABN89580
ID      ABN89580 standard; DNA; 3888 BP.
XX      AC      ABN89580;
XX      DT      06-SEP-2002 (first entry)
XX      DE      Phaseolin promoter-oleosin Trxh-phaseolin terminator DNA SEQ:16.
XX      KW      Multimeric protein; redox protein; thioredoxin; thioredoxin reductase;
XX      oil body; ophthalmological; antidiabetic; cytostatic; antipsoriatic;
XX      vasotropic; vulnereary; antibacterial; immunosuppressive; antitumor;
XX      food product; milk; wheat; oxidative stress; cataract; diabetes;
XX      chronic obstructive pulmonary disease; emphysema; psoriasis; sepsis;
XX      bronchiopulmonary disease; malignancy; reperfusion injury; wound healing;
XX      gastro intestinal bleeding; intestinal bowel disease; ulcer;
XX      gastro oesophageal reflux disease; gene; ds.
XX      Arabidopsis SP.
XX      Phaseolus vulgaris.
XX      WO200250289-A1.
XX      27-JUN-2002.
XX      19-DEC-2001; 2001WO-US50240.
XX      19-DEC-2000; 2000US-0742900.
XX      05-JUL-2001; 2001US-302885P.
XX      04-DEC-2001; 2001US-0006038.
XX      (SEMB-) SEMBIOSYS GENETICS INC.
XX      (SYGN) SYNGENTA PARTICIPATIONS AG.
XX      Van Rooijen G, Deckers H, Helfetz PB, Briggs SP, Dalmia BK;
XX      Del Val G, Zaplachinski S, Moloney M;
XX      WPI: 2002-508806/54.
XX      P-PSDB: ABP60681.
XX      Producing oil body associated with recombinant multimeric protein
XX      complex e.g. redox proteins and immunoglobulins comprises producing
XX      recombinant polypeptides capable of forming the complex in cells
XX      comprising oil bodies.
XX      Example 2: Page 167-168; 362pp; English.
XX      The present invention describes a method (M1) for producing an oil body
XX      associated with a recombinant multimeric protein complex (MPC). M1
XX      comprises producing in a cell comprising oil bodies a first and second
XX      recombinant polypeptide (P1, P2), where P1 is capable of associating
XX      with P2 to form the MPC and associating the complex with an occlusion
XX      body (OB) through an OB-targeting-protein capable of associating with OB
XX      and P1. M1 is useful for producing an oil body associated with a
XX      recombinant MPC. The oil bodies are further formulated for use in the
XX      preparation of a food product such as milk or wheat based food product,
XX      personal care product which reduces the oxidative stress on the surface
XX      area of the human body or used to lighten the skin, or a pharmaceutical
XX      composition used to treat chronic obstructive pulmonary disease (COPD),
XX      cataracts, diabetes, emphysema, bronchiopulmonary disease, psoriasis,
XX      malignancies, reperfusion injury, wound healing, sepsis, gastro
XX      intestinal (GI) bleeding, intestinal bowel disease (IBD), ulcers, GERD

```



PR	21- <u>APR-1999</u>	9905-0130443
PR	23- <u>APR-1999</u>	9905-0130851
PR	23- <u>APR-1999</u>	9905-0130891
PR	28- <u>APR-1999</u>	9905-0131449
PR	30- <u>APR-1999</u>	9905-0132048
PR	30- <u>APR-1999</u>	9905-0132407
PR	04- <u>MAY-1999</u>	9905-0132465
PR	04- <u>MAY-1999</u>	9905-0132485
PR	06- <u>MAY-1999</u>	9905-0132486
PR	07- <u>MAY-1999</u>	9905-0132486
PR	11- <u>MAY-1999</u>	9905-0134263
PR	14- <u>MAY-1999</u>	9905-0134218
PR	14- <u>MAY-1999</u>	9905-0134219
PR	14- <u>MAY-1999</u>	9905-0134321
PR	14- <u>MAY-1999</u>	9905-0134370
PR	18- <u>MAY-1999</u>	9905-0134768
PR	19- <u>MAY-1999</u>	9905-0135491
PR	20- <u>MAY-1999</u>	9905-0135124
PR	21- <u>MAY-1999</u>	9905-0135353
PR	24- <u>MAY-1999</u>	9905-0135629
PR	25- <u>MAY-1999</u>	9905-0136021
PR	27- <u>MAY-1999</u>	9905-0136392
PR	28- <u>MAY-1999</u>	9905-0136787
PR	01- <u>JUN-1999</u>	9905-0137222
PR	03- <u>JUN-1999</u>	9905-0137528
PR	04- <u>JUN-1999</u>	9905-0137502
PR	07- <u>JUN-1999</u>	9905-0137724
PR	08- <u>JUN-1999</u>	9905-0138094
PR	10- <u>JUN-1999</u>	9905-0138540
PR	10- <u>JUN-1999</u>	9905-0138847
PR	14- <u>JUN-1999</u>	9905-0139191
PR	16- <u>JUN-1999</u>	9905-0139452
PR	16- <u>JUN-1999</u>	9905-0139453
PR	17- <u>JUN-1999</u>	9905-0139492
PR	18- <u>JUN-1999</u>	9905-0139455
PR	18- <u>JUN-1999</u>	9905-0139456
PR	18- <u>JUN-1999</u>	9905-0139457
PR	18- <u>JUN-1999</u>	9905-0139458
PR	18- <u>JUN-1999</u>	9905-0139459
PR	18- <u>JUN-1999</u>	9905-0139460
PR	18- <u>JUN-1999</u>	9905-0139461
PR	18- <u>JUN-1999</u>	9905-0139462
PR	18- <u>JUN-1999</u>	9905-0139463
PR	18- <u>JUN-1999</u>	9905-0139750
PR	18- <u>JUN-1999</u>	9905-0139763
PR	21- <u>JUN-1999</u>	9905-0139817
PR	22- <u>JUN-1999</u>	9905-0139819
PR	23- <u>JUN-1999</u>	9905-0140353
PR	23- <u>JUN-1999</u>	9905-0140354
PR	24- <u>JUN-1999</u>	9905-0140655
PR	28- <u>JUN-1999</u>	9905-0140823
PR	29- <u>JUN-1999</u>	9905-0140991
PR	30- <u>JUN-1999</u>	9905-0141287
PR	01- <u>JUL-1999</u>	9905-0141842
PR	01- <u>JUL-1999</u>	9905-0142155
PR	02- <u>JUL-1999</u>	9905-0142554
PR	06- <u>JUL-1999</u>	9905-0142390
PR	08- <u>JUL-1999</u>	9905-0142803
PR	09- <u>JUL-1999</u>	9905-0144920
PR	12- <u>JUL-1999</u>	9905-0145342
PR	13- <u>JUL-1999</u>	9905-0145347
PR	14- <u>JUL-1999</u>	9905-0145362
PR	15- <u>JUL-1999</u>	9905-0144005
PR	16- <u>JUL-1999</u>	9905-0144085
PR	16- <u>JUL-1999</u>	9905-0144086
PR	19- <u>JUL-1999</u>	9905-0144325
PR	19- <u>JUL-1999</u>	9905-0144331
PR	19- <u>JUL-1999</u>	9905-0144332
PR	19- <u>JUL-1999</u>	9905-0144333
PR	19- <u>JUL-1999</u>	9905-0144334
PR	19- <u>JUL-1999</u>	9905-0144335

PR	14-0CT-1999;	990S-0159331.
PR	14-0CT-1999;	990S-0159637.
PR	14-0CT-1999;	990S-0159638.
PR	18-0CT-1999;	990S-0159584.
PR	21-0CT-1999;	990S-0160741.
PR	21-0CT-1999;	990S-0160767.
PR	21-0CT-1999;	990S-0160768.
PR	21-0CT-1999;	990S-0160770.
PR	21-0CT-1999;	990S-0160814.
PR	21-0CT-1999;	990S-0160815.
PR	22-0CT-1999;	990S-0160980.
PR	22-0CT-1999;	990S-0160981.
PR	22-0CT-1999;	990S-0160989.
PR	25-0CT-1999;	990S-0161404.
PR	25-0CT-1999;	990S-0161405.
PR	25-0CT-1999;	990S-0161406.
PR	26-0CT-1999;	990S-0161359.
PR	26-0CT-1999;	990S-0161360.
PR	26-0CT-1999;	990S-0161361.
PR	28-0CT-1999;	990S-0161920.
PR	28-0CT-1999;	990S-0161922.
PR	28-OCT-1999;	990S-0161993.
PR	29-0CT-1999;	990S-0162142.
Alignment Scores:		
Pred. NO.:	7.06e-55	Length: 470
Score:	478.00	Matches: 87
Percent Similarity:	89.29%	Conservative: 13
Best Local Similarity:	77.68%	Mismatches: 12
Query Match:	77.72%	Indels: 0
DB:	21	Gaps: 0
US-09-786-715-6 (1-120) x AAC37781 (1-470)		
Oy	4 SerSerclugludglvglnvallyiesercyshisthrvalgiglutrpasnasplndeu	23
Dd	98 GCTTGGAGACAGACAAGATGATCGCCTCCACACCCTTGACACTGACAGACACTT	157
Oy	24 GlndslglyansgnlsrLysLysLeuileValjaAspherhralaseTrPCysgly	43
Dd	158 CACAAAGCCTAATGAATCAAAAACCTTTGTGTGTGTGAATTACAGCGCTTCGTGTGGA	217
Oy	44 ProccysArghelldlaabrophleuAlaglualeualalylsyphethsrValille	63
Dd	218 CCATGTCGTTCAATCCCTCCATCTTTGCTGATTTGGCTAAGAACAACCTCTTAACGTCCTT	277
Oy	64 PhelnLysValaspPalasplndeuLysSerValsercInasPTripAlalleGluala	83
Dd	278 TTCTCTCAGGTGATCATGATGATTGAAGTGGTGGCAGAGATTTGGCGATACAGCGC	337
Oy	84 MetproThrPheValdPheVallysgluglyThrleuleuasplysValaIglyAlalys	103
Dd	338 ATCCCAACCTTCATCTGTTTTGAAGGAGGAAGATTGTCACAAAGTTGTGGAGCCAG	397
Oy	104 LysASpgluendnglnglnslieglnlyshstVal	115
Dd	398 AAAGATGAGCTTCAGTCCTACCAATTTGCCAAACACTTG	433
RESULT 13		
ID	ABN89586 standard; DNA: 4935 BP.	
XX	ABN89586;	
XX	AC	
XX	DT	
XX	06-SEP-2002 (first entry)	
XX	Promoter-Oleosin thioredoxin reductase-linker-thioredoxin-terminator #33.	
KW	Multimeric protein; redox protein; thioredoxin; thioredoxin reductase;	
KM	oil body; ophthalmological; antidiabetic; cytosolic; antiportatic;	
KW	vasotropy; vulnnerary; antibacterial; immunosuppressive; antidiuric;	
KM	food product; milk; wheat; oxidative stress; cataract; diabetes;	
KW	chronic obstructive pulmonary disease; envenomation; psoriasis; sepsis;	

[illegible][illegible]

DB: 22 Gaps: 0

US-09-786-715-6 (1-120) x AAH87768 (1-509)

QY 1 MetIaGlySerSerGIuGlUgIValIleSerCysHsIthrValGIuGlUtrPAsn 20  
 DB 32 ATGGCTTCGTCGGAATGTGAAGAGAGATGATCGGCTGCCACACCTGATACCTGGAAC 91  
 QY 21 AspGlnLeuGlnLysGlnGlnSerLysLysLeuIleValIAspPheThrAlaSer 40  
 DB 92 GAGCAGCTTCAGAGGGAATGATACAGAGAGTTGGAGTTGCTGATTCATCTGCTTCC 151  
 QY 41 TrpCysGlyProCysArgPheIleAlaProPheLeuAlaGluLeuAlaLysLysPheThr 60  
 DB 152 TGGTGGCGGCTTCGCTTCATCGCCCTTCCTTCACAGAAATGGCCAAAGATTCCT 211  
 QY 61 SerValIlePheLeuLysValAspValAspGluLeuLysSerValSerGlnAspTrpAla 80  
 DB 212 AATGTGACATTTCTCAAGGTGATGTGATGATGAGTGAAGTGGTGGTCTAGTGGGCA 271  
 QY 81 IleGluAlaMetProThrPheValPheValLysGluLysIleGlnLysHisValAlaSerAlaSerAl 100  
 DB 272 GTGGAGGCAATGCCAACCTTCATCTCTCAAGAGAGAGAGATCTTGGACAGCTGCTA 331  
 QY 101 -GlyAlaLysLysAspGluLeuGlnLysIleGlnLysHisValAlaSerAlaSerAl 120  
 DB 332 GGGAGCGCAAGAAAGAGAGAGCTGCGAGCAATATGCTAAGCAGCCTCAACACAGCTACTAG 391  
 QY 120 a 120  
 DB 392 T 392

RESULT 15  
 AA09783 ID AA09783 standard; DNA; 653 BP.  
 XX  
 AC AA09783;  
 DT 24-JUN-1996 (first entry)  
 XX  
 DE Plant SAR gene pl.4.3.  
 XX  
 KW SAR; tobacc; protein-synthesis independent gene; cyclohexamide;  
 KW systemic acquired resistance response; pl.4.3; anti-pathogen;  
 XX  
 OS Nicotiana glauca.  
 XX  
 PN W09519443-A2.  
 XX  
 PD 20-JUL-1995.  
 XX  
 PF 03-JAN-1995; 95MO-IB00002.  
 XX  
 PR 13-JAN-1994; 94US-0181271.  
 XX  
 PA (CIBA ) CIBA GEIGY AG.  
 XX  
 PI Alexander DC, Ryals JA, Uknes SJ, Ward ER;  
 XX  
 DR WPI; 1995-263872/34.  
 XX  
 PT New DNA contg. plant systemic acquired resistance genes - and  
 XX also Arabidopsis gene promoter to control DNA transcription  
 PS Disclosure; Page 61; 85pp; English.  
 CC This sequence represents the DNA sequence of a tobacco protein-synthesis  
 CC independent gene. The gene is involved in the regulation of the  
 CC systemic acquired resistance (SAR) response. This gene is designated  
 CC pl.4.3. This sequence and AA09784 represent SAR genes that are not used  
 CC in the recombinant/chimeric DNA molecules of the invention.

CC AA09784-099790 and AA099806 are SAR genes used in the  
 CC recombinant/chimeric DNA molecules of the invention. The wild type  
 CC genes corresponding to these sequences can all be chemically induced in a  
 CC plant in a protein-synthesis independent manner. SAR genes are involved  
 CC in the protection of plants against pests and disease. These sequences  
 CC were isolated by differential screening of a cDNA library, followed by  
 CC analysis by Northern hybridisation to RNA in the presence and absence of  
 CC cyclohexamide. The genes are used in the creation of transgenic plants.  
 CC Transgenic expression of 2 or more of the recombinant molecules of the  
 CC invention that encode anti-pathogenic proteins provides a synergistic  
 CC increase in plant protection, and may also offer protection against a  
 CC wider range of pathogens.

SQ Sequence 653 BP; 165 A; 114 C; 157 G; 215 T; 2 other;

#### Alignment Scores:

Score: 3.13e-50 Length: 653  
 Percent Similarity: 445.00 Matches: 84  
 Best Local Similarity: 84.87% Conservative: 17  
 Query Match: 70.59% Mismatches: 17  
 DB: 16 Gaps: 0

US-09-786-715-6 (1-120) x AA09783 (1-653)

QY 2 AlaGlySerSerGIuGlUgIValIleSerCysHsIthrValGIuGlUtrPAsn 21  
 DB 63 GCTACTTCATCCGAGGAGGAGCAAGTGTGGCTGCCCAAGAGTTGAGAGATGAGAGAG 122  
 QY 22 GlnLeuGlnLysGlnGlnSerLysLysLeuIleValIAspPheThrAlaSer 41  
 DB 123 TACTTCAAGAAAGCGCTTGAGACTAAGAACTGCTGCTGCTTACTGCTTCAATG 182  
 QY 42 CysGlyProCysArgPheIleAlaProPheLeuAlaGluLeuAlaLysLysPheThrSer 61  
 DB 183 TGGGSCCTTCGCTTTTATTTGCCCCAATTTCTCTGACATGCTTAAGAGATGCCCAT 242  
 QY 62 ValIlePheLeuLysValAspValAspGluLeuLysSerValSer-GlnAspTrpAlaI 81  
 DB 243 GTTATATTCCTCAAGGTTGATGCTGATGACTGAAAGACTGTTACGCGGAGATGAGTGT 302  
 QY 81 eGluAlaMetProThrPheValPheValLysGluLysIleGlnLysHisValAlaSerAlaSer 101  
 DB 303 GAGGCAATGCCAACCTTTGCTTCATTAAGATGAAAGAGTGAAGAGAGAGTGTGG 362  
 QY 101 ValAlaLysLysAspGluLeuGlnLysIleGlnLysHisValAlaSerAlaSer 119  
 DB 363 TGGCAGAGAAAGAGAGGTTGCGAGCAGCATATGATGACATGCTGCTGCTACT 417

Search completed: August 17, 2003, 19:37:14  
 Job time : 173.696 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - nucleic search, using frame\_plus\_p2n-model

Run on: August 17, 2003, 19:18:57 ; Search time 44.472 Seconds  
(without alignments)  
1190.995 Million cell updates/sec

Title: US-09-786-715-6

Perfect score: 615  
Sequence: 1 MAGSSEGGVICHVPEWMN.....GAKKDELOQKIKHVASASA 120

Scoring table: BLOSUM62

Xgapop 10.0 , Xgapext 0.5  
Ygapop 10.0 , Ygapext 0.5  
Fgapop 6.0 , Fgapext 7.0  
Delop 6.0 , Delext 7.0

Searched: 569978 segs, 220691566 residues

Total number of hits satisfying chosen parameters: 1139956

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%  
Listing first 45 summaries

Command line parameters:

-MODE=frame\_plus\_p2n.model -DEV=xlp  
-Q=/cgn2\_1/USPTO.spool/US09786715/runal\_11082003\_150515.6083/app\_query.fasta\_1.1052  
-DB=Issued\_Patents\_NA -QFMT=fastap -SUFFIX=tnl -MINMATCH=0.1 -LOOPEL=0  
-LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blomsun62 -TRANS=human40.cdi  
-LIST=45 -DOCALLIGN=ptc -THR SCORE=ptc -THR MAX=100 -THR MIN=0 -ALIGN=15  
-MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=200000000  
-USER=US09786715.ecgn\_1\_1175.6runat\_11082003\_150515.6083 -MCP=6 -ICPO=3  
-NO\_MMAP -LARGEQUERY -NEG\_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG  
-DEV\_TIMEOUT=120 -WARN\_TIMEOUT=30 -THREDS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6  
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

Issued\_Patents\_NA:\*  
1: /cgn2\_6/ptodata/1/ina/5A.COMB.seq:\*  
2: /cgn2\_6/ptodata/1/ina/5B.COMB.seq:\*  
3: /cgn2\_6/ptodata/1/ina/6A.COMB.seq:\*  
4: /cgn2\_6/ptodata/1/ina/6B.COMB.seq:\*  
5: /cgn2\_6/ptodata/1/ina/PCTUS.COMB.seq:\*  
6: /cgn2\_6/ptodata/1/ina/backfiles1.seq:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	445	72.4	653	1	US-08-181-271A-103
2	445	72.4	653	1	US-08-449-315-103
3	445	72.4	653	1	US-08-444-803-103
4	445	72.4	653	1	US-08-449-043-103
5	445	72.4	653	1	US-08-456-265A-103
6	445	72.4	653	1	US-08-455-416-103
7	445	72.4	653	1	US-08-455-244-103
8	445	72.4	653	1	US-08-454-876-103
9	445	72.4	653	2	US-08-457-364-103
10	445	72.4	653	2	US-08-456-262-103
11	445	72.4	653	2	US-08-456-240-103
12	445	72.4	653	2	US-08-455-736-103

13	445	72.4	653	2	US-08-971-217-103	Sequence 103, App
14	445	72.4	653	3	US-09-350-600-103	Sequence 103, App
15	368	59.8	393	4	US-09-540-014-5	Sequence 5, Appl1
16	367	59.7	382	4	US-09-540-014-3	Sequence 3, Appl1
17	361	58.7	359	4	US-09-540-014-1	Sequence 1, Appl1
18	244.5	39.8	318	4	US-09-313-294A-5868	Sequence 5868, Ap
19	222	36.1	557	4	US-09-404-879A-88	Sequence 88, Appl
20	222	36.1	557	4	US-09-338-933-88	Sequence 88, Appl
21	222	36.1	557	4	US-09-215-681-88	Sequence 88, Appl
22	222	36.1	581	4	US-09-601-144-67	Sequence 87, Appl
23	222	36.1	594	4	US-09-404-879A-87	Sequence 87, Appl
24	222	36.1	594	4	US-09-338-933-87	Sequence 87, Appl
25	222	36.1	554	4	US-09-215-681-87	Sequence 87, Appl
26	222	36.1	601	4	US-09-404-879A-133	Sequence 133, App
27	222	36.1	601	4	US-09-338-933-133	Sequence 133, App
28	222	36.1	601	4	US-09-215-681-133	Sequence 133, App
29	222	36.1	624	3	US-09-385-982-526	Sequence 526, App
30	222	36.1	630	3	US-08-180-371-5	Sequence 5, Appl1
31	222	36.1	630	3	US-08-180-371-17	Sequence 17, Appl
32	222	36.1	630	5	PCT-US92-05707-5	Sequence 5, Appl1
33	211	34.3	605	3	US-09-385-982-485	Sequence 485, App
34	203	33.0	631	3	US-09-385-982-174	Sequence 174, App
35	199	32.4	278	4	US-09-313-294A-5840	Sequence 5840, App
36	195.5	31.8	914	1	US-08-386-729A-9	Sequence 9, Appl1
37	182.5	29.7	339	4	US-09-107-532A-3085	Sequence 3085, Ap
38	178.5	29.0	336	4	US-09-134-001C-1022	Sequence 1022, Ap
39	173	28.1	564	2	US-08-775-878-2	Sequence 2, Appl1
40	171	27.8	7096	4	US-09-221-017B-373	Sequence 373, App
41	170.5	27.7	1230025	4	US-09-198-452A-1	Sequence 1, Appl1
42	163.5	26.6	4403765	3	US-09-103-840A-2	Sequence 2, Appl1
43	163.5	26.6	4411529	3	US-09-103-840A-1	Sequence 1, Appl1
44	159	25.9	1830121	4	US-09-557-884-1	Sequence 1, Appl1
45	159	25.9	1830121	4	US-09-643-990A-1	Sequence 1, Appl1

#### ALIGNMENTS

RESULT 1  
US-08-181-271A-103  
; Sequence 103, Application US/08181271A  
; Patent No. 5614395  
; GENERAL INFORMATION:  
; APPLICANT: Ryals, John A.  
; APPLICANT: Alexander, Danny C.  
; APPLICANT: Beck, James J.  
; APPLICANT: Duesing, John H.  
; APPLICANT: Friedlich, Leslie B.  
; APPLICANT: Goodman, Robert M.  
; APPLICANT: Harms, Christian  
; APPLICANT: Meins, Jr., Frederick  
; APPLICANT: Montoya, Alice  
; APPLICANT: Moyer, Mary B.  
; APPLICANT: Neuhaus, Jean-Marc  
; APPLICANT: Payne, George B.  
; APPLICANT: Sperison, Christoph  
; APPLICANT: Stinson, Jeffrey R.  
; APPLICANT: Uknes, Scott J.  
; APPLICANT: Ward, Eric R.  
; TITLE OF INVENTION: CHEMICALLY REGULATABLE AND ANTI-PATHOGENIC  
; TITLE OF INVENTION: DNA SEQUENCES AND USES THEREOF  
; NUMBER OF SEQUENCES: 106  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: CIBA-GEIGY Corporation  
; STREET: 7 Skyline Drive  
; CITY: Hawthorne  
; STATE: New York  
; COUNTRY: USA  
; ZIP: 10532  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patentin Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/181,271A  
FILING DATE: 13-JAN-94  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/093,301  
FILING DATE: 16-JUL-1993  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/937,197  
FILING DATE: 6-NOV-1992  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/678,378  
FILING DATE: 1-APR-1991  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/305,566  
FILING DATE: 6-FEB-1989  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/165,667  
FILING DATE: 8-MAR-1988  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/042,847  
FILING DATE: 6-APR-1993  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/632,441  
FILING DATE: 21-DEC-1990  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/425,504  
FILING DATE: 20-OCT-1989  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/848,506  
FILING DATE: 6-MAR-1992  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/768,122  
FILING DATE: 27-SEP-1991  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/580,431  
FILING DATE: 7-SEP-1990  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/368,672  
FILING DATE: 20-JUN-1989  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/329,018  
FILING DATE: 24-MAR-1989  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/045,957  
FILING DATE: 12-APR-1993  
ATTORNEY/AGENT INFORMATION:  
NAME: Elmer, James Scott  
REGISTRATION NUMBER: 36,129  
REFERENCE/DOCKET NUMBER: S-19825/P1/GC 1727  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (919)541-8614  
TELEFAX: (919)541-8689  
INFORMATION FOR SEQ ID NO: 103:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 653 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: DNA (genomic)  
US-08-181-271A-103

Alignment Scores:  
Pred. No.: 1,686-54 Length: 653  
Score: 445.00 Matches: 84  
Percent Similarity: 84.87% Conservative: 17  
Best Local Similarity: 70.59% Mismatches: 17  
Query Match: 72.36% Indels: 1  
DB: 1 Gaps: 0

US-09-786-715-6 (1-120) x US-08-181-271A-103 (1-653)

QY 2 AlaGlySerSerGluGluGlyGlnValIleSerCysHisThrValGluGluTrpAsnAsp 21  
Db 63 GCTACTGATCCGAGAGGAGGACAAAGTTCGCGTCCACAAAGGTTGAGAAATGCAACGAG 122  
QY 22 GlnLeuGlnGlyAspGluSerLysLysLeuIleValAspPheThrAlaSerTrp 41  
Db 123 TACTTCAGAAAGCGGTGAGACTAAGAACTGGGTGGTGGATTACTTACGCTTCACAGG 182  
QY 42 CysGlyProCysArgPheIleAlaProPheLeuAlaGluLeuAlaLysLysPheThrSer 61  
Db 183 TGGCGSCCTCCGCTTTATTTGCCCCATTTCTGTCACATTGCAATAAAGATGCCCAT 242  
QY 62 ValIlePheLeuLysValAspValAspGluLeuLysSerValSer-GlnAspTrpAlaIle 81  
Db 243 GTTATATTCCTCAAGGTTGATGTGATCACTGAAGACTGTTTCACCGGAATGAGTGT 302  
QY 81 egluAlaMetProThrPheValPheValLysGluGlyThrLeuLeuAspLysValValG1 101  
Db 303 GGAGGCAATGCCCACTTTTGTCTTCATTAAAGATGGAAGAAAGAGTGCACAGATTGTTGG 362  
QY 101 YAlaLysLysAspGluLeuGlnGlnLysIleGlnLysHisValAlaSerAlaSer 119  
Db 363 TGCCAAAGAAAGAGGAGTTGCAGACACCATAGTGAGAGCATGCTGCTGCTACT 417

RESULT 2  
US-08-449-315-103  
Sequence 103, Application US/08449315  
Patent No. 5650505  
GENERAL INFORMATION:  
APPLICANT: Ryals, John A.  
APPLICANT: Alexander, Danny C.  
APPLICANT: Beck, James J.  
APPLICANT: Duesing, John H.  
APPLICANT: Friedrich, Leslie B.  
APPLICANT: Goodman, Robert M.  
APPLICANT: Harms, Christian  
APPLICANT: Meins, Jr., Frederick  
APPLICANT: Montoya, Alice  
APPLICANT: Moyer, Mary B.  
APPLICANT: Neuhaus, Jean-Marc  
APPLICANT: Payne, George B.  
APPLICANT: Sperison, Christoph  
APPLICANT: Stinson, Jeffrey R.  
APPLICANT: Uknes, Scott J.  
APPLICANT: Ward, Eric R.  
TITLE OF INVENTION: Chemically Regulatable And Anti-Pathogenic  
TITLE OF INVENTION: DNA SEQUENCES AND USES THEREOF  
NUMBER OF SEQUENCES: 106  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: CIBA-GEIGY Corporation  
STREET: 7 Skyline Drive  
CITY: Hawthorne  
STATE: New York  
COUNTRY: USA  
ZIP: 10532  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/449,315  
FILING DATE: 24-MAY-1995  
CLASSIFICATION: 800  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/181,271  
FILING DATE: 13-JAN-94  
APPLICATION NUMBER: US 08/093,301  
FILING DATE: 16-JUL-1993  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/937,197  
FILING DATE: 6-NOV-1992

PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/678,378  
FILING DATE: 1-APR-1991  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/305,566  
FILING DATE: 6-FEB-1989  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/165,667  
FILING DATE: 8-MAR-1988  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/042,847  
FILING DATE: 6-APR-1993  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/632,441  
FILING DATE: 21-DEC-1990  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/425,504  
FILING DATE: 20-OCT-1989  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/848,506  
FILING DATE: 6-MAR-1992  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/768,122  
FILING DATE: 27-SEP-1991  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/580,431  
FILING DATE: 7-SEP-1990  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/368,672  
FILING DATE: 20-JUN-1989  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/329,018  
FILING DATE: 24-MAR-1989  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/045,957  
FILING DATE: 12-APR-1993  
ATTORNEY/AGENT INFORMATION:  
NAME: Elmer, James Scott  
REGISTRATION NUMBER: 36,129  
REFERENCE/DOCKET NUMBER: S-19825/PL/CGC 1727  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (919)541-8614  
TELEFAX: (919)541-8689  
INFORMATION FOR SEQ ID NO: 103:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 653 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: DNA (genomic)  
US-08-449-315-103  
Alignment Scores:  
Pred. No.: 1,68e-54 Length: 653  
Score: 445.00 Matches: 84  
Percent Similarity: 84.87% Conservative: 17  
Best Local Similarity: 70.59% Mismatches: 17  
Query Match: 72.36% Indels: 1  
DB: 1 Gaps: 0  
US-09-786-715-6 (1-120) x US-08-449-315-103 (1-653)  
QY 2 AlaGlySerSerGluGluGlyValIleSerGlyThrValGluGluTrpAsp 21  
DB 63 GCTACTTATCCGAGAGGACCAAGTGTTCGCGTCCACCAAGTTGAGAGAAAGAGAG 122  
QY 22 GlnLeuGlnLysGlyAsnGlySerLysLysLeuIleValAlaAspPheThrAlaSerTrp 41  
DB 123 TACTTCAGAGAAAGCGGTGAGAGCTAAGAACTGCGTGCATTTACTCTCTCATGG 182  
QY 42 CysGlyProCysArgPheIleAlaProPheLeuAlaGluLeuAlaLysLysPheThrSer 61  
DB 183 TCGGSGCTTCGCGTTTATTCGCCCAATTCCTGCTGACATTCGTAAGAAGATGCCCAT 242

QY 62 ValIlePheLeuLysValAspValAlaAspGluLeuLysSerValSer-GlnAspTrpAlaIle 81  
DB 243 GTTATATTCCTCAAGGTGATGTGATGACAGAGACTGTTTCACGGGAAATGAGAGTGT 302  
QY 81 eGluAlaMetProThrPheValPheValLysGluGlyThrLeuLeuAspLysValValAla 101  
DB 303 GGAGCAATGCCAAGCTTTGTTCTTCATTAAGATGAAAGAAAGATGACAGAGTGTCTGG 362  
QY 101 ValAlaLysLysAspGluLeuGlnGlnLysIleGlnLysHisValAlaSerAlaSer 119  
DB 363 TCCCAAGAAAGAGAGAGTTCACAGACACCACTAGTGAAGCATGCTGCTGCTACT 417  
RESULT 3  
US-08-444-803-103  
Sequence 103, Application US/08444803  
Patent No. 5654414  
GENERAL INFORMATION:  
APPLICANT: Ryals, John A.  
APPLICANT: Alexander, Danny C.  
APPLICANT: Beck, James J.  
APPLICANT: Duestling, John H.  
APPLICANT: Friedrich, Leslie B.  
APPLICANT: Goodman, Robert M.  
APPLICANT: Harms, Christian  
APPLICANT: Meins, Jr., Frederick  
APPLICANT: Montoya, Alice  
APPLICANT: Moyer, Mary B.  
APPLICANT: Neuhaus, Jean-Marc  
APPLICANT: Payne, George B.  
APPLICANT: Sperison, Christoph  
APPLICANT: Stinson, Jeffrey R.  
APPLICANT: Uknes, Scott J.  
APPLICANT: Ward, Eric R.  
APPLICANT: Williams, SheriLeca C.  
TITLE OF INVENTION: CHEMICALLY REGULATABLE AND ANTI-PATHOGENIC  
TITLE OF INVENTION: DNA SEQUENCES AND USES THEREOF  
NUMBER OF SEQUENCES: 106  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: CIBA-GEIGY Corporation  
STREET: 7 Skyline Drive  
CITY: Hawthorne  
STATE: New York  
COUNTRY: USA  
ZIP: 10532  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/444,803  
FILING DATE: 19-MAY-1995  
CLASSIFICATION: 536  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/181,271  
FILING DATE: 13-JAN-94  
APPLICATION NUMBER: US 08/093,301  
FILING DATE: 16-JUL-1993  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/937,197  
FILING DATE: 6-NOV-1992  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/678,378  
FILING DATE: 1-APR-1991  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/305,566  
FILING DATE: 6-FEB-1989  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/165,667  
FILING DATE: 8-MAR-1988  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/042,847

FILING DATE: 6-APR-1993  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/632,441  
FILING DATE: 21-DEC-1990  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/425,504  
FILING DATE: 20-OCT-1989  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/848,506  
FILING DATE: 6-MAR-1992  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/768,122  
FILING DATE: 27-SEP-1991  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/580,431  
FILING DATE: 7-SEP-1990  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/368,672  
FILING DATE: 20-JUN-1989  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/329,018  
FILING DATE: 24-MAR-1989  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/045,957  
FILING DATE: 12-APR-1993  
ATTORNEY/AGENT INFORMATION:  
NAME: Elmer, James Scott  
REGISTRATION NUMBER: 36,129  
REFERENCE/DOCKET NUMBER: S-19825/PL/CGC 1727  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (919)541-8614  
TELEFAX: (919)541-8689  
INFORMATION FOR SEQ ID NO: 103:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 653 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: DNA (genomic)  
US-08-444-803-103

Alignment Scores:  
Pred. No.: 1.68e-54 Length: 653  
Score: 445.00 Matches: 84  
Percent Similarity: 84.87% Conservative: 17  
Best Local Similarity: 70.59% Mismatches: 17  
Query Match: 72.36% Indels: 1  
DB: 1 Gaps: 0

US-09-786-715-6 (1-120) x US-08-444-803-103 (1-653)

QY 2 AlaglyserSerSglugllyglValIleSerCysHsthrValGluGluTrpAsnasp 21  
DB 63 GCTACTTCATCCAGGAGGAGCAAGTGTGGCTGCACAGGTGAGCAATGGAACGAG 122  
QY 22 GlnLeuGlnLysGlnSnglSerLysLysLeuIleValValAspPheThrAlaSerTyr 41  
DB 123 TACTTCAGAGAAAGCGCTGAGACTGAAGAACTGCTGCTCATTTACTGCTTCATGG 182  
QY 42 CysGlyProCysArgPheIleAlaProPheLeuAlaGluLeuAlaLysLysPheThrSer 61  
DB 183 TGGGSGCTTCGCTTTATTGCGCCCAATTCCTGCTGACATTCCTAAGAAAGATGCCCAT 242  
QY 62 ValIlePheLeuLysValAspValAspGluLeuLysSerValSer-GlnAspTrpAlaIle 81  
DB 243 GTATATTCCTCAGAGTGTGATGATGAACCTGAAGACTTTTCAGCGGGAATGCACTGT 302  
QY 81 eGluAlaMetProThrPheValAlaPheValLysGlnGlyThrLeuLeuAspLysValAlaGln 101  
DB 303 GGAGGCAATGCCAAGCTTTCTTCATTAAGATGAGAAAGAGTGAAGAGTGTG 362  
QY 101 yAlaLysLysAspGlnLeuGlnInLysIleGlnLysHisValAlaSerLaser 119  
|||||

Db 363 TGCCAGAAAGAGAGTTCAGACGACACCACTAGTGAAGCANTGCTCTCTCTACT 417

RESULT 4  
US-08-449-043-103  
Sequence 103, Application US/08449043  
Patent No. 5689044  
GENERAL INFORMATION:  
APPLICANT: Ryals, John A.  
APPLICANT: Alexander, Danny C.  
APPLICANT: Beck, James J.  
APPLICANT: Duesing, John H.  
APPLICANT: Friedrich, Leslie B.  
APPLICANT: Goodman, Robert M.  
APPLICANT: Harms, Christian  
APPLICANT: Meins, Jr., Frederick  
APPLICANT: Montoya, Alice  
APPLICANT: Moyer, Mary B.  
APPLICANT: Neunhaus, Jean-Marc  
APPLICANT: Payne, George B.  
APPLICANT: Sperison, Christoph  
APPLICANT: Stinson, Jeffrey R.  
APPLICANT: Uknes, Scott J.  
APPLICANT: Ward, Eric R.  
TITLE OF INVENTION: Williams, Shericca C.  
TITLE OF INVENTION: CHEMICALLY REGULATABLE AND ANTI-PATHOGENIC  
TITLE OF INVENTION: DNA SEQUENCES AND USES THEREOF  
NUMBER OF SEQUENCES: 106  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: CIBA-GEIGY Corporation  
STREET: 7 Skyline Drive  
City: Hawthorne  
STATE: New York  
COUNTRY: USA  
ZIP: 10532

COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/449,043  
FILING DATE: 24-MAY-1995  
CLASSIFICATION: 800  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/181,271  
FILING DATE: 13-JAN-94  
APPLICATION NUMBER: US 08/093,301  
FILING DATE: 16-JUL-1993  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/937,197  
FILING DATE: 6-NOV-1992  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/678,378  
FILING DATE: 1-APR-1991  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/305,566  
FILING DATE: 6-FEB-1989  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/165,667  
FILING DATE: 8-MAR-1988  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/042,847  
FILING DATE: 6-APR-1993  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/632,441  
FILING DATE: 21-DEC-1990  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/425,504  
FILING DATE: 20-OCT-1989  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/848,506  
FILING DATE: 6-MAR-1992  
PRIOR APPLICATION DATA:





LENGTH: 653 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: DNA (genomic)  
US-08-456-265A-103

Alignment Scores:  
Pred. No.: 1,68e-54 Length: 653  
Score: 445.00 Matches: 84  
Percent Similarity: 84.87% Conservative: 17  
Best Local Similarity: 70.59% Mismatches: 17  
Query Match: 72.36% Indels: 1  
DB: 1 Gaps: 0

US-09-786-715-6 (1-120) x US-08-456-265A-103 (1-653)

QY 2 AAlaGlySerSerGluGluGluGluValIleSerCysHisThrValGluGluTrpAsnAsp 21  
DB 63 GCTACTTCATCCGAGGAGGAGCAAGTTCGCGCCACAGGTTGAGAAATGAGACGAG 122  
QY 22 GlnLeuGlnYsgIYsAngIuSerLysLysLeuIleValValAspPheThrAlaSerTrp 41  
DB 123 TACTTCAGAAAGCGCTTGAGACTAAGAACTGGTGGTGCATTTTACTGCTTCATGG 182  
QY 42 CysGlyProCysArGpHeIleAlaProPheLeuAlaGluLeuAlaLysLysPheThrSer 61  
DB 183 TGGGSCCTTGCGCTTTATTTGCCCCCATCTCTGCATGTGTAAGAAGATGCCCAT 242  
QY 62 ValIlePheLeuLysValAspValAspGluLeuLysSerValSer-GlnAspTrpAlaIle 81  
DB 243 GTTATATTCCTCAAGGTGTGATGTGATGAAGTGAAGCTGTGAGCGGCAATGAGAGTCT 302  
QY 81 eGluAlaMetProThrPheValPheValIleValIleGlyThrLeuLeuAspLysValValG 101  
DB 303 GGAGGCAATGCCACTTTTGTCTCATTAAGATGGAAGAAAGATGGACAGATGTTGG 362  
QY 101 yAlaLysLysAspGluLeuGlnGlnLysIleGlnLysHisValAlaSerAlaSer 119  
DB 363 TGCAGAAAGAGAGGATGTGCAGCAGACCATAGTGAACATGCTCTGCTACT 417

RESULT 6  
US-08-455-416-103  
Sequence 103, Application US/08455416  
Patent No. 5777200  
GENERAL INFORMATION:  
APPLICANT: Ryals, John A.  
APPLICANT: Alexander, Danny C.  
APPLICANT: Beck, James J.  
APPLICANT: Duesing, John H.  
APPLICANT: Friedrich, Leslie B.  
APPLICANT: Goodman, Robert M.  
APPLICANT: Harms, Christian  
APPLICANT: Meins, Jr., Frederick  
APPLICANT: Montoya, Alice  
APPLICANT: Moyer, Mary B.  
APPLICANT: Neunhaus, Jean-Marc  
APPLICANT: Payne, George B.  
APPLICANT: Sperison, Christoph  
APPLICANT: Stinson, Jeffrey R.  
APPLICANT: Uknes, Scott J.  
APPLICANT: Ward, Eric R.  
APPLICANT: Williams, Shericea C.  
TITLE OF INVENTION: CHEMICALLY REGULATABLE AND ANTI-PATHOGENIC  
TITLE OF INVENTION: DNA SEQUENCES AND USES THEREOF  
NUMBER OF SEQUENCES: 106  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: CIBA-GEIGY Corporation  
STREET: 7 Skyline Drive  
CITY: Hawthorne  
STATE: New York  
COUNTRY: USA  
ZIP: 10532

COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/455,416  
FILING DATE: 31-MAY-1995  
CLASSIFICATION: 800  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/181,271  
FILING DATE: 13-JAN-94  
APPLICATION NUMBER: US 08/093,301  
FILING DATE: 16-JUL-1993  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/937,197  
FILING DATE: 6-NOV-1992  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/678,378  
FILING DATE: 1-APR-1991  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/305,566  
FILING DATE: 6-FEB-1989  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/165,667  
FILING DATE: 8-MAR-1988  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/042,847  
FILING DATE: 6-APR-1993  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/632,441  
FILING DATE: 21-DEC-1990  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/425,504  
FILING DATE: 20-OCT-1989  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/848,506  
FILING DATE: 6-MAR-1992  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/768,122  
FILING DATE: 27-SEP-1991  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/580,431  
FILING DATE: 7-SEP-1990  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/368,672  
FILING DATE: 20-JUN-1989  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/329,018  
FILING DATE: 24-MAR-1989  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/045,957  
FILING DATE: 12-APR-1993  
ATTORNEY/AGENT INFORMATION:  
NAME: Elmer, James Scott  
REGISTRATION NUMBER: 36,129  
REFERENCE/DOCKET NUMBER: S-19825/P1/CGC 1727  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (919)541-8614  
TELEFAX: (919)541-8689  
INFORMATION FOR SEQ ID NO: 103:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 653 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: DNA (genomic)  
US-08-455-416-103

Alignment Scores:  
Pred. No.: 1,68e-54 Length: 653  
Score: 445.00 Matches: 84  
Percent Similarity: 84.87% Conservative: 17



Db 123 TACTTCAAGAAAGCGTTGACACTAGAACTGGTGGTGCATTTTACTGCTTCATGCG 182  
Qy 42 CysglyProCysatgphelilealProPheleuAlaGluleuAlaLysPheThSer 61  
Db 183 TCGGSCCTTGCCTTTTATTGCCCCCAATTCCTGTCATGCTTAAGAAATGCCCAT 242  
Qy 62 ValIlePheLeuLysValAspValAspGluLeuLysSerValSer-GlnAspTTPalaIl 81  
Db 243 GTTATATTCCTCAAGGTTGATGATGATGAACTGTTTCAAGCGGAAATGAGAGTGT 302  
Qy 81 eGIAlaMetProThrPheValPheValLysGluGlyThrLeuLysPlysValValGI 101  
Db 303 GGAGCGCATGCCAATTTTGTCTTCATTAAGATGAAAGAAAGTGCAGACAGTGTGG 362  
Qy 101 YAlaLysAspGluLeuGlnLysIleGlnLysHisValAlaSerAlaSer 119  
Db 363 TGCACAAAGAGAGGATTGCAGACCAATAGTGAAGCATGCTCTCTCTACT 417

## RESULT 8

US-08-454-876-103  
Sequence 103: Application US/08454876  
Patent No. 5804693

## GENERAL INFORMATION:

APPLICANT: Ryals, John A.  
APPLICANT: Alexander, Danny C.  
APPLICANT: Beck, James J.  
APPLICANT: Duesing, John H.  
APPLICANT: Friedrich, Leslie B.  
APPLICANT: Goodman, Robert M.  
APPLICANT: Harms, Christian  
APPLICANT: Melns, Jr., Frederick  
APPLICANT: Monloya, Alice  
APPLICANT: Moyer, Mary B.  
APPLICANT: Neuhaus, Jean-Marc  
APPLICANT: Payne, George B.  
APPLICANT: Sperison, Christoph  
APPLICANT: Stinson, Jeffrey R.  
APPLICANT: Uknes, Scott J.  
APPLICANT: Ward, Eric R.  
APPLICANT: Williams, Shericea C.  
TITLE OF INVENTION: CHEMICALLY REGULATABLE AND ANTI-PATHOGENIC  
NUMBER OF SEQUENCES: 106  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: CIBA-GEIGY Corporation  
STREET: 7 Skyline Drive  
CITY: Hawthorne  
STATE: New York  
COUNTRY: USA  
ZIP: 10532  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/454,876  
FILING DATE: 31-MAY-1995  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/181,271  
FILING DATE: 13-JAN-94  
APPLICATION NUMBER: US 08/093,301  
FILING DATE: 16-JUL-1993  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/937,197  
FILING DATE: 6-NOV-1992  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/678,378  
FILING DATE: 1-APR-1991  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/305,566

Db 123 TACTTCAAGAAAGCGTTGACACTAGAACTGGTGGTGCATTTTACTGCTTCATGCG 182  
Qy 42 CysglyProCysatgphelilealProPheleuAlaGluleuAlaLysPheThSer 61  
Db 183 TCGGSCCTTGCCTTTTATTGCCCCCAATTCCTGTCATGCTTAAGAAATGCCCAT 242  
Qy 62 ValIlePheLeuLysValAspValAspGluLeuLysSerValSer-GlnAspTTPalaIl 81  
Db 243 GTTATATTCCTCAAGGTTGATGATGATGAACTGTTTCAAGCGGAAATGAGAGTGT 302  
Qy 81 eGIAlaMetProThrPheValPheValLysGluGlyThrLeuLysPlysValValGI 101  
Db 303 GGAGCGCATGCCAATTTTGTCTTCATTAAGATGAAAGAAAGTGCAGACAGTGTGG 362  
Qy 101 YAlaLysAspGluLeuGlnLysIleGlnLysHisValAlaSerAlaSer 119  
Db 363 TGCACAAAGAGAGGATTGCAGACCAATAGTGAAGCATGCTCTCTCTACT 417

FILING DATE: 6-FEB-1989  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/165,667  
FILING DATE: 8-MAR-1988  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/042,847  
FILING DATE: 6-APR-1993  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/632,441  
FILING DATE: 21-DEC-1990  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/425,504  
FILING DATE: 20-OCT-1989  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/848,506  
FILING DATE: 6-MAR-1992  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/768,122  
FILING DATE: 27-SEP-1991  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/580,431  
FILING DATE: 7-SEP-1990  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/368,672  
FILING DATE: 20-JUN-1989  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/329,018  
FILING DATE: 24-MAR-1989  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/045,957  
FILING DATE: 12-APR-1993  
ATTORNEY/AGENT INFORMATION:  
NAME: Elmer, James Scott  
REGISTRATION NUMBER: 36,129  
REFERENCE/DOCKET NUMBER: S-19825/P1/CGC 1727  
TELEPHONE: (919)541-8614  
TELEFAX: (919)541-8689  
INFORMATION FOR SEQ ID NO: 103:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 653 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: DNA (genomic)  
US-08-454-876-103

Alignment Scores:  
Pred. No.: 1,68e-54 Length: 653  
Score: 445.00 Matches: 84  
Percent Similarity: 84.87% Conservative: 17  
Best Local Similarity: 70.59% Mismatches: 17  
Query Match: 72.36% Indels: 1  
DB: 1 Gaps: 0

US-09-786-715-6 (1-120) x US-08-454-876-103 (1-653)

Qy 2 AlaGlySerSerGluGluGlyGlnValIleSerCysHisThrValGluGluTrpAsp 21  
Db 63 GCTACTTCATCCGAGGAGGAGGAGCACTGTCCGCTGCCCAAGATTGAGATGAGAG 122  
Qy 22 GlnLeuGlnLysGlyAsnGluSerLysLysLeuIleValValAspPheThrAlaSerTrp 41  
Db 123 TACTTCAAGAAAGCGTTGACACTAGAACTGGTGGTGCATTTTACTGCTTCATGCG 182  
Qy 42 CysglyProCysatgphelilealProPheleuAlaGluleuAlaLysPheThSer 61  
Db 183 TCGGSCCTTGCCTTTTATTGCCCCCAATTCCTGTCATGCTTAAGAAATGCCCAT 242  
Qy 62 ValIlePheLeuLysValAspValAspGluLeuLysSerValSer-GlnAspTTPalaIl 81  
Db 243 GTTATATTCCTCAAGGTTGATGATGATGAACTGTTTCAAGCGGAAATGAGAGTGT 302

```
QY      81  eGIuAlaMeTProThrPheValAlPheValIysGluGlyThrLeuAspLysValValG1 101
      :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db      303 GGAGGCAATGCCACTTTTGTCTTCATTAAGATGAAAGAAAGAGACAGAGTGTGTG 362
QY      101 yAlAlsYsAspGluLeuGlnGlnLysIleGlnLysValAlAserAlAser 119
      :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db      363 TGCCAAAGAAAGAGAGTTCACACACATAGTGAAGCATGCTGCTCCTACT 417

RESULT 9
US-08-457-364-103
; Sequence 103, Application US/08457364
; Patent No. 5847258
; GENERAL INFORMATION:
; APPLICANT: Ryals, John A.
; APPLICANT: Alexander, Danny C.
; APPLICANT: Beck, James J.
; APPLICANT: Duesling, John H.
; APPLICANT: Friedlich, Leslie B.
; APPLICANT: Goodman, Robert M.
; APPLICANT: Harms, Christian
; APPLICANT: Meins, Jr., Frederick
; APPLICANT: Montoya, Alice
; APPLICANT: Meyer, Mary B.
; APPLICANT: Neuhaus, Jean-Marc
; APPLICANT: Payne, George B.
; APPLICANT: Sperison, Christoph
; APPLICANT: Stinson, Jeffrey R.
; APPLICANT: Uknes, Scott J.
; APPLICANT: Ward, Eric R.
; APPLICANT: Williams, Sherocca C.
; TITLE OF INVENTION: CHEMICALLY REGULATABLE AND ANTI-PATHOGENIC
; TITLE OF INVENTION: DNA SEQUENCES AND USES THEREOF
; NUMBER OF SEQUENCES: 106
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: CIBA-GEIGY Corporation
; STREET: 7 Skyline Drive
; CITY: Hawthorne
; STATE: New York
; COUNTRY: USA
; ZIP: 10532
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/457,364
; FILING DATE: 31-MAY-1995
; CLASSIFICATION: 800
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/181,271
; FILING DATE: 13-JAN-94
; APPLICATION NUMBER: US 08/093,301
; FILING DATE: 16-JUL-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/937,197
; FILING DATE: 6-NOV-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/678,378
; FILING DATE: 1-APR-1991
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/305,566
; FILING DATE: 6-FEB-1989
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/165,667
; FILING DATE: 8-MAR-1988
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/042,847
; FILING DATE: 6-APR-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/632,441
; FILING DATE: 21-DEC-1990
; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: US 07/425,504
; FILING DATE: 20-OCT-1989
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/848,506
; FILING DATE: 6-MAR-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/768,122
; FILING DATE: 27-SEP-1991
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/580,431
; FILING DATE: 7-SEP-1990
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/368,672
; FILING DATE: 20-JUN-1989
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/329,018
; FILING DATE: 24-MAR-1989
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/045,957
; FILING DATE: 12-APR-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: Elmer, James Scott
; REGISTRATION NUMBER: 36,129
; REFERENCE/DOCKET NUMBER: S-19825/P1/CCG 1727
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (919)541-8614
; TELEFAX: (919)541-8689
; INFORMATION FOR SEQ ID NO: 103:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 653 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; US-08-457-364-103

Alignment Scores:
Pred. No.: 1,68e-54 Length: 653
Score: 445.00 Matches: 84
Percent Similarity: 84.878 Conservative: 17
Best Local Similarity: 70.594 Mismatches: 1
Query Match: 72.364 Indels: 1
Gaps: 0

US-09-786-715-6 (1-120) x US-08-457-364-103 (1-653)
QY      2  ALaGlySerSerGlnGlnGlnGlnValIleSerCysHisThrValGlnGlnTrpAsnAsp 21
      :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db      63  GCTACTTCATCCGAGAGGAGGACAGTGTGGCTGCCACACAGTTGAGCAATGGAACGAG 122
QY      22  GlnLeuGlnLysGlnGlnGlnSerLysLysLeuIleValValAspPheThrAlaSerTrp 41
      :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db      123 TACTTCACAGAAAGGCGGTGACACTAAGAAACGTGGTGCTGATTCTCTCATCATG 182
QY      42  CysGlyProCysArgPheIleLeuAlaProPheLeuAlaGlnLeuAlaLysSerPheThrSer 61
      :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db      183 TCGGSGCTTCCGCTTTTATTTGCCCCAATTTCTGTCGACATTGCTTAGAGATGCCCAT 242
QY      62  ValIlePheLeuLysValAspValAspGlnLeuLysSerValSer-GlnAspTrpAlaIle 81
      :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db      243 GTTATATTCCTCAAGGTGATGCTGATGAACAGAGCTGTTACAGCGGAAATGAGAGTGT 302
QY      81  eGIuAlaMeTProThrPheValAlPheValIysGluGlyThrLeuAspLysValValG1 101
      :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db      303 GGAGGCAATGCCACTTTTGTCTTCATTAAGATGAAAGAAAGAGACAGAGTGTGTG 362
QY      101 yAlAlsYsAspGluLeuGlnGlnLysIleGlnLysValAlAserAlAser 119
      :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db      363 TGCCAAAGAAAGAGAGTTCACACACATAGTGAAGCATGCTGCTCCTACT 417

RESULT 10
US-08-456-262-103
; Sequence 103, Application US/08456262
```

Patent No. 5851766  
GENERAL INFORMATION:  
APPLICANT: Ryals, John A.  
APPLICANT: Alexander, Danny C.  
APPLICANT: Beck, James J.  
APPLICANT: Duesing, John H.  
APPLICANT: Friedrich, Leslie B.  
APPLICANT: Goodman, Robert M.  
APPLICANT: Harms, Christian  
APPLICANT: Meins, Jr., Frederick  
APPLICANT: Montoya, Alice  
APPLICANT: Moyer, Mary B.  
APPLICANT: Neunus, Jean-Marc  
APPLICANT: Payne, George B.  
APPLICANT: Sperison, Christoph  
APPLICANT: Stinson, Jeffrey R.  
APPLICANT: Uknes, Scott J.  
APPLICANT: Ward, Eric R.  
TITLE OF INVENTION: CHEMICALLY REGULATABLE AND ANTI-PATHOGENIC  
TITLE OF INVENTION: DNA SEQUENCES AND USES THEREOF  
NUMBER OF SEQUENCES: 106  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: CIBA-GEIGY Corporation  
STREET: 7 Skyline Drive  
CITY: Hawthorne  
STATE: New York  
COUNTRY: USA  
ZIP: 10532  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/456,262  
FILING DATE: 31-MAY-1995  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/181,271  
FILING DATE: 13-JAN-94  
APPLICATION NUMBER: US 08/093,301  
FILING DATE: 16-JUL-1993  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/937,197  
FILING DATE: 6-NOV-1992  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/678,378  
FILING DATE: 1-APR-1991  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/305,566  
FILING DATE: 6-FEB-1989  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/165,667  
FILING DATE: 8-MAR-1988  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/042,847  
FILING DATE: 6-APR-1993  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/632,441  
FILING DATE: 21-DEC-1990  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/425,504  
FILING DATE: 20-OCT-1989  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/848,506  
FILING DATE: 6-MAR-1992  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/768,122  
FILING DATE: 27-SEP-1991  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/580,431  
FILING DATE: 7-SEP-1990

PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/368,672  
FILING DATE: 20-JUN-1989  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/329,018  
FILING DATE: 24-MAR-1989  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/045,957  
FILING DATE: 12-APR-1993  
ATTORNEY/AGENT INFORMATION:  
NAME: Elmer, James Scott  
REGISTRATION NUMBER: 36,129  
REFERENCE/DOCKET NUMBER: S-19825/PL/CGC 1727  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (919)541-8614  
TELEFAX: (919)541-8689  
INFORMATION FOR SRO ID NO: 103:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 653 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: DNA (genomic)  
US-08-456-262-103  
Alignment Scores:  
Pred. No.: 1.68e-54 Length: 653  
Score: 445.00 Matches: 84  
Percent Similarity: 84.87% Conservative: 17  
Best Local Similarity: 70.59% Mismatches: 17  
Query Match: 72.36% Indels: 1  
Gaps: 0  
US-09-786-715-6 (1-120) x US-08-456-262-103 (1-653)  
QY 2 AlaglysersegluglglglnvaliliserCyshisrhyValgluglutrpsanasp 21  
Db 63 GCTACTTCATCCAGGAGGAGCAAGTTCGCTGCCACCAAGTTGAGGATGGAAGCAG 122  
QY 22 GlnleuglnysglnglsnglsrlyslsleuileValalasphehrilaserTrp 41  
Db 123 TACTTCAGAAAGCCGTTGAGCTAGAAACTGTGTGCTTCATTTTACTGCTTCAATGG 182  
QY 42 CysglpProCysarqpherlealeaProphelueuAgluAgluAgluAgluAgluAglu 61  
Db 183 TGGGSGCCTTGCCTTTATTTGCCCCAATTCCTGCTGACATTCCTAAGAAAGATGCCCAT 242  
QY 62 ValillepheleulysvalaspValaspGluLeulyssevalser-Glnasptpalail 81  
Db 243 GTTATATTCCTCAAGGTTGATGATGAGACTGAGACTGTTTCAGCGGAATGAGGTGT 302  
QY 81 egluAlaMetProThrPheValPheVallyslgslgylThrLeuLeuaspLysValValgl 101  
Db 303 GGAAGCAATGCCAAGCTTTTGTCTTCATTAAAGATGGAAGAAGATGAGACAGATTGTTGG 362  
QY 101 yAlaLysAspGluLeuglnGlnlylleglnLysHisValAlaSerAlaser 119  
Db 363 TGGCAAGAAAGAGAGAGTTCAGACAGACATAGGAAGCATGCGTCTCTACT 417  
RESULT 11  
US-08-456-240-103  
Sequence 103, Application US/08456240  
Patent No. 5856154  
GENERAL INFORMATION:  
APPLICANT: Ryals, John A.  
APPLICANT: Alexander, Danny C.  
APPLICANT: Beck, James J.  
APPLICANT: Duesing, John H.  
APPLICANT: Friedrich, Leslie B.  
APPLICANT: Goodman, Robert M.  
APPLICANT: Harms, Christian  
APPLICANT: Meins, Jr., Frederick  
APPLICANT: Montoya, Alice

```

REGISTRATION NUMBER: 36,129
REFERENCE/DOCKET NUMBER: S-19825/P1/CGC 1727
TELECOMMUNICATION INFORMATION:
TELEPHONE: (919)541-8614
TELEFAX: (919)541-8689
INFORMATION FOR SEQ ID NO: 103:
SEQUENCE CHARACTERISTICS:
LENGTH: 653 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
US-08-456-240-103

Alignment Scores:
Pred. No.: 1,68e-54 Length: 653
Score: 445.00 Matches: 84
Percent Similarity: 84.87% Conservative: 17
Best Local Similarity: 70.59% Mismatches: 17
Query Match: 72.36% Indels: 1
DB: Gaps: 0

US-09-786-715-6 (1-120) x US-08-456-240-103 (1-653)
QY      2  AAlAglySerSeGluGlUGlYglValIleesCySHstHrValGIuglUTrpaSnp 21
        GGTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTT
Db      63  GTACTCTTCGCGAGGAGGACAGTGTCGGCTGCCCAAGGTGGGAATGGAAAGCAG 122
QY     22  GlNLeuGlInLysGlYAncIUserLySLySLeuLIleValIAsppHeThAlasErTri 41
        :::::||||| ||:::||||| :::::||||| :::::|||||
Db    123  TACTTCACAAGAAGGCGGTGAGACTAAGAACTGCGTGGTGGCATTTTACTGCTTATTGG 182
QY     42  CysGlyProCYsArgPheIIleAlaPrObheuleAlaGUleuAlaLysLysPheThrSer 61
        |||||GGGSCCTGGCGCTTTATTGCCCAATTCCTGCTGACATTCCTAAGAAGATGCCCAT 242
Db    183  TCGGSGCCTGGCGCTTTATTGCCCAATTCCTGCTGACATTCCTAAGAAGATGCCCAT 242
QY     62  ValIIIPheLeuLYsValASpYalAspLUleuLYSSerValSer-GlnaspTrrAlaIl 81
        |||||eGuaIametProThrPheValIPheValIleSgLUlyThrLeuLeuAspLYsValIAl 101
Db    303  GAGGCGAATGCCAACCTTTTGCTCTTCATTAAAGATGGAAAAGAGTGCACAGATTGTGG 362
QY     101  yAlaLYsLYsAspGIuLeuGInGLnLYsIIeGInLYsHISValAIaseralaser 119
        |||||:::||||| :::::||||| :::::||||| ||| ||:::
Db    363  TGCCAAGAAGAAGGAGGTTCAGACACACATGTAAGACATGCTGCTCTGCTACT 417

RESULT 12
US-08-453-736-103
; Sequence 103, Application US/08455736
; Patent No. 5880328
GENERAL INFORMATION:
APPLICANT: Ryals, John A.
APPLICANT: Alexander, Danny C..
APPLICANT: Beck, James J.
APPLICANT: Duesling, John H.
APPLICANT: Friedrich, Leslie B.
APPLICANT: Goodman, Robert M.
APPLICANT: Harms, Christian
APPLICANT: Weins, Jr., Frederick
APPLICANT: Montoya, Alice
APPLICANT: Moyer, Mary B.
APPLICANT: Neuhaus, Jean-Marc
APPLICANT: Payne, George B.
APPLICANT: Sperison, Christoph
APPLICANT: Stinson, Jeffrey R.
APPLICANT: Uknes, Scott J.
APPLICANT: Ward, Eric R.
TITLE OF INVENTION: CHEMICALLY REGULATABLE AND ANTI-PATHOGENIC
NUMBER OF SEQUENCES: 106

```







FILING DATE: 6-APR-1993  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/632,441  
FILING DATE: 21-DEC-1990  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/425,504  
FILING DATE: 20-OCT-1989  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/848,506  
FILING DATE: 6-MAR-1992  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/768,122  
FILING DATE: 27-SEP-1991  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/580,431  
FILING DATE: 7-SEP-1990  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/368,672  
FILING DATE: 20-JUN-1989  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/329,018  
FILING DATE: 24-MAR-1989  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/381,443  
FILING DATE: 18-JUL-1989  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/353,312  
FILING DATE: 17-MAY-1989  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/226,303  
FILING DATE: 29-JUL-1988  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/045,957  
FILING DATE: 12-APR-1993  
ATTORNEY/AGENT INFORMATION:  
NAME: Mels, J. Timothy  
REGISTRATION NUMBER: 38,241  
REFERENCE/DOCKET NUMBER: S-198250  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (919)541-8587  
TELEFAX: (919)541-8689  
INFORMATION FOR SEQ ID NO: 103:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 653 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: DNA (genomic)  
US-09-350-600-103

Alignment Scores:  
Pred. No.: 1.68e-54 Length: 653  
Score: 445.00 Matches: 84  
Percent Similarity: 84.87% Conservative: 17  
Best Local Similarity: 70.59% Mismatches: 17  
Query Match: 72.36% Indels: 1  
Gaps: 0

US-09-786-715-6 (1-120) x US-09-350-600-103 (1-653)

QY 2 AAlGlyserSerGluGluGluGluValIleSerCysHisThrValGluGluTrpAsp 21  
Db 63 GCTACTTCATCCAGGAGGCAAGTGTGGCTGCCACCAAGTGTGAGATGGACGAC 122  
QY 22 GlnLeuGlnLysGlnLysLeuSerLysLysLeuIleValValAspPheThrAlaSerTrp 41  
Db 123 TACTTCAGAAAGCGCTGTGAGACTAGAAACTGGTGTGCTTCATTCTTCATGG 182  
QY 42 CysGlyProCysArgPheIleAlaProPheLeuAlaGluValLysLysPheThrSer 61  
Db 183 TGGGSCCTTGCGCTTTATTGGCCCAATCTTGTCACATGTGTAAGAGATGCCCAT 242  
QY 62 ValIlePheLeuLysValAspValAspGluLeuLysSerValSer-GlnAspTrpAlaIle 81

|||||  
Db 243 GTTATATTCCTCAAGGTGTGATGTGATGAAGACTGTTTCACGGCAATGAGTGT 302  
QY 81 eGluAlaMetProThrPheValIleValLysGluGlyThrLeuLysPheValValGlu 101  
Db 303 GGAGGCAATGCCAATTTTGTCTTCATTAAGATGAAAGAGATGGACAGATGTTGG 362  
QY 101 ValAlaLysAspGluLeuGlnGlnLysIleGluLysHisValAlaSerAlaSer 119  
Db 363 TGCCAGAAAGAGAGAGATTTGCCAGCAGACCATATGTAAGACATGCTCTGCTACT 417

RESULT 15  
US-09-540-014-5  
Sequence 5, Application US/09540014  
Patent No. 6380372  
GENERAL INFORMATION:  
APPLICANT: Cho, Myeong-Je  
APPLICANT: Del Val, Greg  
APPLICANT: Caliau, Maxime  
APPLICANT: Lemaux, Peggy G.  
APPLICANT: Buchanan, Bob B.  
TITLE OF INVENTION: Barley Gene for Thioresdoxin and  
TITLE OF INVENTION: NADP-Thioresdoxin Reductase  
FILE REFERENCE: 2001-0701.30  
CURRENT FILING DATE: US/09/540,014  
PRIOR APPLICATION NUMBER: 2000-03-31  
PRIOR FILING DATE: 1999-03-31  
PRIOR APPLICATION NUMBER: US 60/169,162  
PRIOR FILING DATE: 1999-12-06  
PRIOR APPLICATION NUMBER: US 60/177,740  
PRIOR FILING DATE: 2000-01-21  
PRIOR APPLICATION NUMBER: US 60/177,739  
PRIOR FILING DATE: 2000-01-21  
NUMBER OF SEQ ID NOS: 51  
SOFTWARE: FastSeq for Windows Version 4.0  
SEQ ID NO 5  
LENGTH: 393  
TYPE: DNA  
ORGANISM: Artificial Sequence  
FEATURE:  
OTHER INFORMATION: wheat thioresdoxin h cDNA  
US-09-540-014-5

Alignment Scores:  
Pred. No.: 8.65e-44 Length: 393  
Score: 368.00 Matches: 63  
Percent Similarity: 81.82% Conservative: 27  
Best Local Similarity: 57.27% Mismatches: 20  
Query Match: 59.84% Indels: 0  
Gaps: 0

US-09-786-715-6 (1-120) x US-09-540-014-5 (1-393)

QY 8 GlyGlnValIleSerCysHisThrValGluGluTrpAsnAspGlnLeuGlnLysGlyAsn 27  
Db 61 GGGAGGTGATCTCCGTCACAGCCTGAGCAGTGGACCATGAGATGAGAGGCCAAC 120  
QY 28 GluSerLysLysLeuIleValValAspPheThrAlaSerTrpCysGlyProCysArgPhe 47  
Db 121 GCCGCAAGAGAGCTGGTGGTGTGATGTGATGACTTACTGACATGATGGACCATGCCGATT 180  
QY 48 IleAlaProPheLeuAlaGluLeuAlaLysLysPheThrSerValIlePheLeuVal 67  
Db 181 ATGGCTCCATTTTGTGATCTGCCAAGATGCCACCTGCTGTTTCTCAAGGTC 240  
QY 68 AspValAspGluLeuLysSerValSerGlnAspTrpAlaIleGluAlaMetProThrPhe 87  
Db 241 GACGTTGATGAATGAAAGCCCATGTCGACCAATTGACGCTCGAGGCCATGCAACCTTC 300  
QY 88 ValPheValLysGluGluThrLeuLeuAspLysValValGluLysLysAspGluLeu 107  
Db 301 CTGTTCAATGAAGAGAGAGACGTCAGACAGGAGGTTGTGGAGCTATCAAGAGAGAGCTG 360



OY 108 GInGInLysIleGInLysHISValAlaSer 117  
|||::|||  
Db 361 ACGACCAAGGTGGGCTCCACGGCGCTGCC 390

Search completed: August 17, 2003, 21:45:22  
Job time : 53.472 secs